Chapter

1

Overview of Health Informatics & Healthcare Data

By Bambang Parmanto, PhD

Learning Objectives

- Define health informatics and how it relates to the emerging field of data science
- Discuss the health data life cycle
- Describe the drivers behind health informatics
- Examine the unique challenges of health informatics
- Interpret the health informatics core competencies
- Articulate the transformation of data into knowledge and wisdom

KEY TERMS

Data
Data acumen
Data literacy
DIKW pyramid
Health data science
Health informatics
Information
Knowledge
Semantic gap
Wisdom

Introduction

Health informatics (HI) is becoming an increasingly important component of healthcare due to the digitization of healthcare processes. The field emerged as a discipline in the 1960s, at which time it played only a niche role. However, current challenges faced by the healthcare system have led to it taking on a more central role. One specific challenge to healthcare that the HI field can address is the digitization of the healthcare process, which requires health information systems to capture, store, retrieve, and exchange information related to every aspect of healthcare. Health informatics plays a key role in the development, implementation, and evaluation of such systems. Moreover, digitization of the healthcare process also has generated large amounts of complex data that need to be transformed into information and knowledge. Health informatics is key to enabling this growth of medical knowledge, which has resulted in novel treatments and interventions. These new treatments and interventions, in turn, require additions to the informatics infrastructure to foster their implementation. Health informatics implements and evaluates the informatics infrastructure to support novel treatments and interventions.

Health informatics is an interdisciplinary field, involving fields such as computer and information sciences, medicine and health sciences, and management sciences. Many definitions of HI can be found in published literature and all over the Internet, but it is challenging to find a succinct definition. However, it is important to provide a definition that serves as a common frame of reference for the discussions in this book. Such a definition should clearly delineate which areas are and are not part of the field as well as define the relationship between HI and other related fields. The definition that we use is drawn from the conceptual model from the 2016 ACMI Winter Symposium (Payne et al. 2018). The adapted model is presented in figure 1.1.

Health informatics is the field concerned with health data acquisition, management, representation, and transformation into knowledge that can subsequently be translated into novel health interventions or new healthcare processes. Figure 1.1 illustrates the complete cycle of HI that includes data transformation to knowledge on the right side of the diagram and the process developing, implementing, and evaluating tools for health intervention on the left. The right side of the cycle represents the growing field of health data science. **Health data science** focuses on transforming health data into knowledge that can be used for making a decision. Both health data science and HI transform data from the world into knowledge (Payne et al. 2018). This has led to the perspective that HI and health data science are the same thing (Ohno-Machado 2013; Fridsma 2018; Payne et al. 2018). As you can see in figure 1.1, HI is a broader field that includes health data science. Health informatics does not stop with the discovery of new knowledge; it also uses knowledge to create interventions and tools that affect the world.

Health informatics harnesses the power of digital technology and applies methods from computer science, information science, and cognitive science for the acquisition, storage, retrieval, use, and sharing of health data, information, and knowledge. The evolution of the HI field follows the pattern of evolution in the field of computing and information technology. In the early stages of HI's evolution, it was primarily focused on the basic recording of medical and healthcare data and organizing those data so that they could be retrieved easily to support daily operations. The HI field was dominated by activities related to the development, implementation, and evaluation of new tools and software systems to support healthcare delivery. Looking at figure 1.1, it focused on the left side of the diagram from the "implement" to "healthcare process" steps.

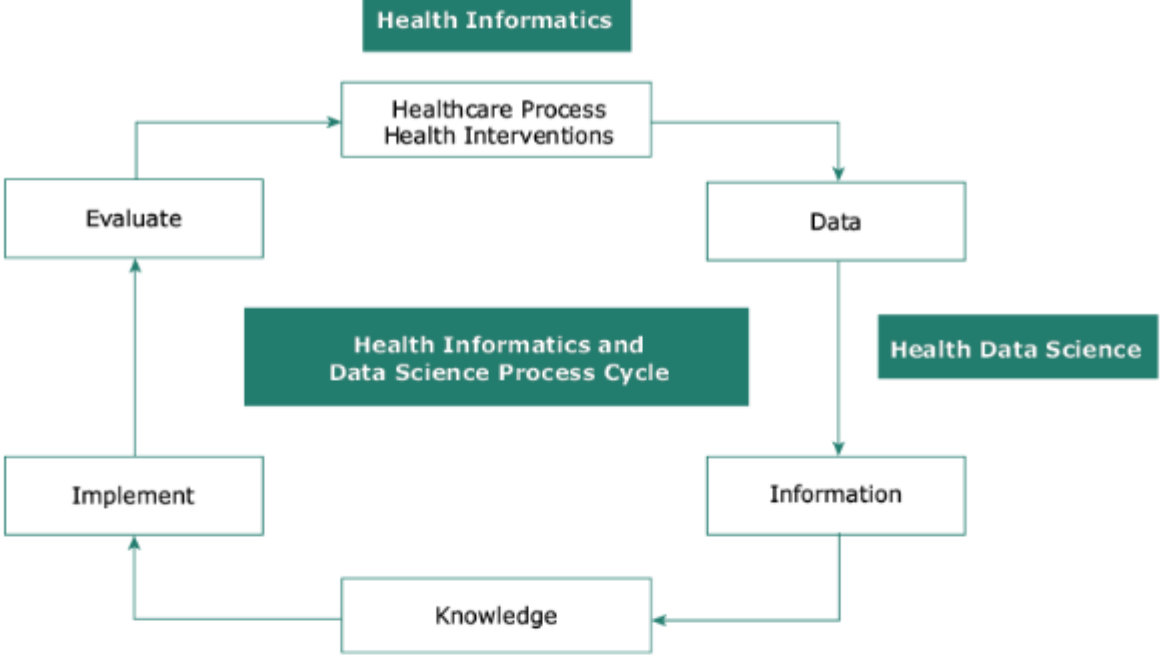


Figure 1.1. Relationship between health informatics and data science in data and knowledge life cycle

▶ Expand/collapse Extended Description

However, as software systems and tools such as electronic health record (EHR) systems became universally adopted and most aspects of healthcare became digital, and the amount of data accumulated from healthcare process grew, the focus of HI evolved. Today, not only the volume of data, but also the rate of increase in the volume of data continues to accelerate. These developments have led to a shift in emphasis of HI from the left side to the right side of the cycle shown in figure 1.1.

Overview of the Book Using the Health Data Life Cycle

This chapter provides an overview of HI as a discipline using a different view of the health data life cycle as illustrated in figure 1.2. This chapter also provides an overview of the contents of the book. The purpose is to give a general idea about how health data are generated, recorded, managed, transferred, and transformed into information and knowledge for decision-making.

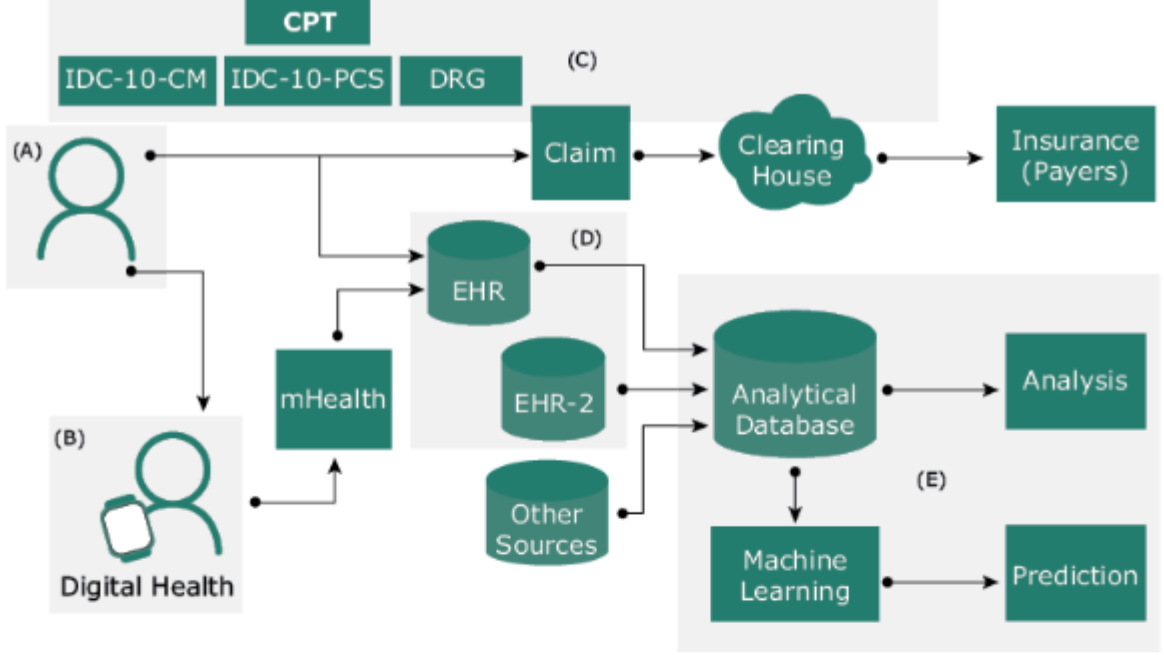


Figure 1.2. Health data cycle and information system infrastructure

▶ Expand/collapse Extended Description

All of the steps in the health data life cycle have to be managed using health information systems. As illustrated in figure 1.1 on the left side of the cycle, new interventions or healthcare processes need to be implemented and supported by health information systems. Implementation and evaluation of health information systems are important roles of HI. In fact, as mentioned earlier, in the early stages of HI as a field, implementation and evaluation of health information systems were the main component of HI. Implementing healthcare information systems is an important component of HI. This topic is covered in Chapter 4, Implementing Healthcare Information Systems.

The health data life cycle starts with data acquisition when a patient visits a healthcare provider. In figure 1.2, this start of the health data lifecycle is labeled as (A). More health data will be generated in the future by individuals as they go about their daily lives when they use wearable devices or smartphones to record their physical activities (B). This topic is covered in Chapter 14, Technology for Digital Health: Telemedicine and Mobile Health Systems. One of the unique aspects of data in HI is that it involves standards, including vocabulary standards that assign meaning to the data (C in figure 1.2). This is important when we discuss transformation of data into information, which is covered in Chapter 3, Health Care Standards, Classification Systems, and Healthcare Information Exchange. Part of the healthcare information will be sent to entities such as insurance companies for payment and reimbursement, while other parts will be sent to a healthcare information exchange for population health data aggregation and analyses. Data exchange and interoperability are essential in HI because health data need to be transferred from one entity to another as part of the healthcare process (reimbursement, reporting, monitoring, and so on). Because of this, standards are a very important topic in HI.

Patient data and healthcare transactions are usually recorded using an information system that is called an EHR system. This is illustrated in (D) in figure 1.2. EHR is a general term to refer to any information system that is used to manage patients health data, ranging from billing to medical history. The topics related to EHR and clinical information systems are covered more extensively in Chapter 2, Electronic Health

Records and Clinical Information Systems. Health data management, including EHR, involves the process of storing and organizing data, which is usually handled using database systems. Relational database systems have been used in the healthcare industry for several decades, and new types of database systems (for example, NoSQL and object oriented) that can store more complex data involving imaging and physiological sensors are increasingly used for modern healthcare applications. Most healthcare data transactions are recorded using database management systems. Understanding how data are organized in the database is an important concept discussed in Chapter 5, Understanding Databases. The transactional data must be aggregated and combined before being analyzed.

HI is increasingly becoming a science that transforms data into knowledge moving beyond the basic traditional role of acquiring, storing, and retrieving health data. As more and more data continue to accumulate, the challenge has shifted from recording transactional data to analyzing data and turning it into information that can be used for making decisions (E). The challenge lies in how to organize data so that they can be analyzed using methods such as machine learning. This topic is discussed in Chapter 8, Introduction to Data Science. One of the core issues in data science is how to integrate data from various sources and to aggregate, clean, and match data from disparate sources.

Once the data are integrated and wrangled, they are ready for analysis. The analyses can be done using statistical tools or machine learning. One of the most powerful tools in turning data into information and knowledge is machine learning. This topic is discussed in Chapter 9, Data Analytics and Machine Learning). The main purpose of data analysis and machine learning prediction is to support decision-making. The relationship between HI and decision-making is discussed in Chapter 7, Introduction to Research.

As illustrated in the health data life cycle in figure 1.2, patient data in a healthcare system move from one department to another, and often from one institution to another for payment or reporting. Patient data can also be combined with other related data that come from other sources or be combined with other population data using location or other demographic characteristics. To make a decision and during analytical phases, data need to be linked and aggregated. All of these data movements and transformations need to be managed to ensure data integrity. The process of managing availability, integrity, and security of the data is called data governance, which is discussed in Chapter 12, Data Governance.

Protecting the privacy and confidentiality of patients is one of the main issues in HI. Protection of patient privacy and confidentiality involves the technical aspects related to health information system security and the other aspects governing patient privacy. The privacy of health information systems is discussed in Chapter 10, Privacy for Health Information. Chapter 11, Security for Health Information, covers security. Health data and new methods that are effective in addressing big data such as machine learning also engender new challenges such as susceptibility to implicit biases that can potentially have serious implications in decision-making. Ethics touches every component and step in the health data life cycle. Chapter 13, Ethics, discusses issues of ethics in HI.

Check Your Understanding 1.1

1. Which one of these activities is not part of health informatics?
   a. Transforming of data into knowledge
   b. Developing tools to support translation of knowledge into new intervention
   c. Data acquisition, management, and representation
   d. Collecting financial and administrative data

2. What is the relationship between health informatics and health data science?
    a. Health informatics is a subset of health data science
    b. Health data science is a subset of health informatics
    c. Health informatics intersects with health data science
    d. Health informatics is separate from health data science
3. The following is not an example of data capture in health data life cycle:
    a. Patient encounter with clinician
    b. Remote monitoring of data of a patient
    c. Wearable devices
    d. Billing data
4. Healthcare standards and classification systems are not important for:
    a. Healthcare data exchange
    b. Reimbursement
    c. Electronic health record
    d. Machine learning

Forces That Shape the Health Informatics Field

The HI field continually changes with the development of newer and more advanced technology and applications. Many underlying forces are propelling HI to the forefront of the health care industry in this century. It is useful to understand the forces that move the HI field to see how and why changes take place. The forces that move the field can be classified into three main categories: methods of data generation, increasing volume of data, and methods of dealing with data accumulation.

New Methods of Data Generation: Digitization of Healthcare

In the healthcare industry, the American Recovery and Reinvestment Act of 2009 (ARRA 2009) was an important milestone that shaped the HI field. This legislation offered monetary incentives to healthcare providers qualified as meaningful users of certified EHR systems. This incentive continues to motivate organizations to implement EHR systems (Washington et al. 2017). To qualify as a meaningful user, the provider must utilize a system that provides automated information that is retrievable in a timely manner by various users in different places, and that ensures patient confidentiality. At the same time, healthcare consumers are more proactive than ever about the quality and cost of their healthcare and demand information at their fingertips.

New Methods in Data Generation: New Data Captures

The digitization of health data goes beyond EHR systems. The advent of wearable devices connected to smartphones allows health data to be captured continuously, generating health data points with every heartbeat. The health data can include physiological data as well as behavioral data. Problems such as depression and anxiety can be assessed or reported right when individuals are experiencing them, rather than waiting for a meeting with a clinician later.

Ever Increasing Volume of Data

The digitization of the healthcare process as well as the expansion of data captures resulted in a huge volume of data. The volume of data in the healthcare industry is projected to grow faster than that in other sectors. The compound annual growth rate of data for healthcare will reach 36 percent in 2025, which is faster than that of financial

services, manufacturing, and media and entertainment (Kent 2018). The role of HI will become more important, as the healthcare industry generates approximately 30 percent of the world's data volume today. This growth is fueled by wearable medical devices and by the digitization of health records, including medical imaging data. The convergence of wearable consumer devices and medical technology has created potential future growth.

New Methods of Analyzing Data

With the amount of data being generated and the continuously increasing complexity of the data, it is impossible to analyze the data without the help of new methods and new technologies. The invention of new methods, such as deep learning, and new technologies, such as big data management and analytics, makes it possible to take advantage of the large and complex data set and turn it into information and knowledge that can be used for decision-making.

Challenges Unique to Health Data

Informatics is a general field that is used in many industries, but one might wonder why health and healthcare needs HI whereas other fields do not have their own specialized field of informatics. It is because health data are unique compared to data in other industries. The challenge with health data is not only its volume, but also its complexity. Unlike the data of any other industry, health data are unique for a number of reasons (LeSueur 2017).

- The contents and types of health data available make it complex, and it will only get more complex as more types become available. The contents of health data can range from demographic and financial to radiological imaging data. It consists of different data types: numbers, images, videos, and text. The data will get more complex in the future as it will include data from wearable sensors to genomic data.
- The data are a mix of both structured and unstructured. Structured data are data organized into specific fields with each field having a defined purpose. Demographic data and financial information are examples of structured data. Unstructured data do not have any predefined data models. Examples of unstructured data include physicians' notes, images, videos, transcripts and recordings, emails, and portal chats. The unique nature of health data is that unstructured data can make up more than 80 percent of data within a healthcare organization. Unstructured data are more difficult to manage and to analyze.
- The nature of health data means a **semantic gap** exists. The word "semantic" here is used to refer to as the meaning of the term or data item. The problem, then, is that the difference between data and information (meaning) in health data is large.
- Much of the data are in multiple places. As illustrated in the data life cycle figure, health data can be generated in various places during clinical encounters and the data are stored in multiple places. Consider that an individual might change healthcare providers or be treated in different hospitals. Even in the same healthcare organization, a patient might visit a primary care that uses an EHR system to record data, while the images are stored in a radiology information system.

Check Your Understanding 1.2

1. Which one is not among the forces that shape health informatics?
   a. New methods of data captures

b. New reimbursement method
c. New methods of analyzing data
d. Digitization of healthcare

2. Large volume of data can be an opportunity. Which one of the statements below is not a reason why it can be an opportunity.
   a. New methods of analyzing data such as deep learning can turn large amounts of data into knowledge
   b. Managing the ever-increasing volume of data can be challenging
   c. Information from a large volume of data can make decision-making more accurate
   d. Large volume of data can potentially provide information that would otherwise not be available

3. Why are health data more complex than data from other industries? Which one of these statements is not the reason?
   a. Health data may contain clinical, demographic, and financial data
   b. Health data can be a combination of images, text, and videos
   c. Health data are not as complex as data from the financial industry
   d. Health data can contain structured and unstructured data
4. Health data are usually stored in multiple places. Which one of these scenarios more likely contributes to the situation?
   a. Patient can go to different specialists
   b. Patient can go to different hospitals
   c. Patient does not have health insurance
   d. Patient can be prescribed by primary care to get images from radiology

---

From Data to Knowledge

Transformation of information is central to health informatics. In figure 1.1, the right side of the HI cycle is a process that transforms data to knowledge. It is therefore important to bring up the context of information hierarchy, commonly known as the **DIKW pyramid**, which describes the relationship between data, information, knowledge, and wisdom first introduced by Ackoff (1989). Figure 1.3 illustrates the DIKW pyramid and how data are transformed into knowledge and wisdom. Through a process of filtration, aggregation, reduction, and transformation, data, which are in plentiful supply, move upward to become information, knowledge, and eventually wisdom (Sardar 2020). For practical purposes, we usually focus on the transformation of data into knowledge that can then be used to make a decision with regard to healthcare processes or new interventions. The ability to use actionable knowledge for data-driven decisions is called wisdom. There is also a perspective that refers to actionable knowledge as intelligence instead of wisdom.

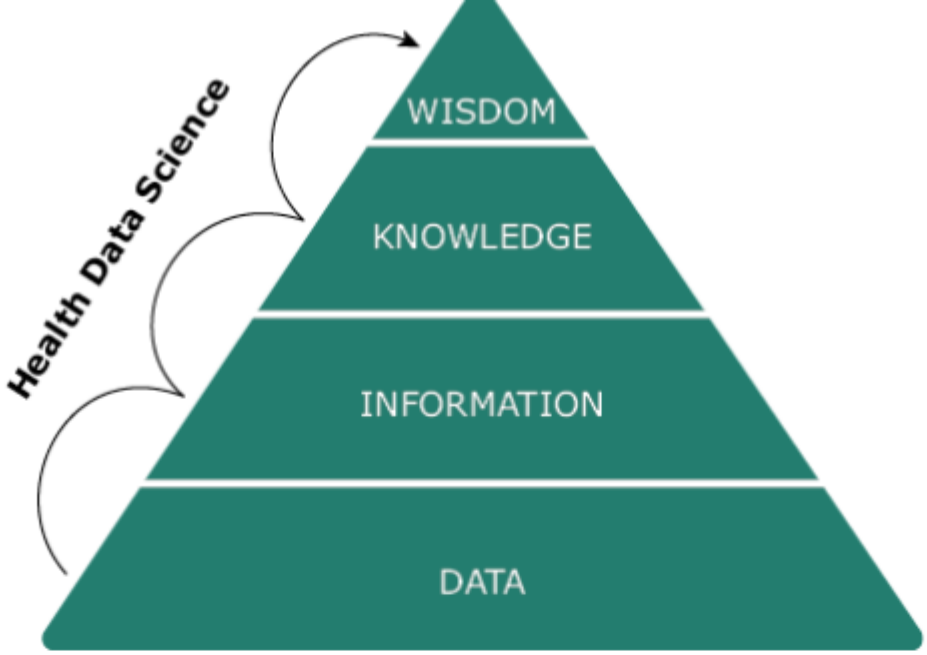


Figure 1.3. DIKW pyramid

▶ Expand/collapse Extended Description

Data Science and DIKW Pyramid: Relationship between Data

Table 1.1 provides a description and defining features of each step in the DIKW pyramid. ***Data*** are raw facts taken from observation. They are discrete, unorganized, raw, and objective facts. In healthcare, data are any raw facts that are recorded in any transaction. An example of a raw fact is a baby's birth weight. By themselves, data are essentially useless because they are isolated facts that do not mean much. Data can be transformed into ***information*** by providing context. An example of context with baby's birth weight is related data such as mother's age, race, residence, and the term of the birth. This context transforms data into information. In other words, information is what the data tell us: who the baby is, where the baby was born, and so on. Data become useful and relevant information when we begin to organize and give it structure. Information adds context to a representation and tells the recipient something that was not known before.

When information is processed or framed by previous experiences and by connecting it with other related information over time, then it is transformed into ***knowledge***. When we evaluate organized information and analyze the relationship between a baby's birth weight and term of labor, we might find that low birth weight is associated with preterm labor, the race of the mother, or the mother's socioeconomic background. Knowledge becomes ***wisdom*** when it becomes integrated into practice, thus allowing the knowledge to be applied in different situations in a measurable way to make intelligent data-driven decisions on a consistent basis. An example of wisdom related to the preceding scenario on birth weight data would be a public health intervention that targets mothers from socioeconomic backgrounds that are associated with low birth weight.

Table 1.1. Examples and defining features of elements in the DIKW pyramid

Level	Descriptions	Defining Feature	Example

Wisdom	Ability to apply knowledge to solve problems	Intelligent, data-driven decisions	Public health programs targeting mothers from certain socioeconomic backgrounds
Knowledge	Group of interrelated pieces of information	Connections and patterns	Low birth weight babies are associated with mother's socioeconomic status
Information	Data combined with context to produce meaning	Interpretations: contextualized, organized, and structured	Baby's birth weight and mother's age, race, and residence
Data	Collection of raw facts	Observations: discrete, unorganized raw objective facts	Baby's birth weight

Skills Required for Health Informatics and Health Data Science Professionals

The HI profession has experienced significant change in the past decade due to the growth in big data and the emergence of the data science field. This growth has brought with it an increased focus on health data science within the HI field. As a result, the skills that are necessary to master in the HI field are a combination of the traditional HI skills and the new and growing types of skills required to work in the field of health data science.

Figure 1.1 at the beginning of this chapter illustrates the relationship between HI as a broader field where health data science is part of it. On the left side of the HI cycle are the more traditional areas of HI. The core skills that are needed for implementation and evaluation of a health information system include health information system implementation, project management, organizational theory, human-computer interaction, usability studies, and health IT infrastructure. The core skills of HI are drawn from the three fields that make up an interdisciplinary field of HI: information sciences, health sciences, and management sciences. Table 1.2 presents some of the core skills of HI. The skills listed are not exhaustive; they are meant to act as examples of HI core skills.

In figure 1.1 on the right side of the HI cycle is the growing area of health data science. As discussed earlier in the section "Forces That Shape the Health Informatics Field," digitization of healthcare and new ways of capturing health data have resulted in a massive volume of data. Health data science can be defined as the field that is concerned with learning from health data to develop health interventions or to improve healthcare processes. Advances in information technology, computation, and machine learning make it possible to access, integrate, and analyze massive amounts of data to transform them into knowledge and use them for data-driven decision-making. The field of HI is increasingly becoming a health data science.

Table 1.2. Interdisciplinary skills in health informatics

Health Sciences	Information Sciences	Management Sciences	Intersections
Healthcare systems (U.S. or a country healthcare system)	Security and privacy of health information	Project management	Digital health interventions
Healthcare standards and classification systems	Database design and management	Quality improvements	Legal and ethical issues in health IT

Health vocabulary and terminology	Human-computer interactions	Financial management	Health data governance

Table 1.3. Data acumen areas

Data Acumen Areas	HI Field
Foundations: statistics and computing	Information sciences and statistics
Data management: database design and retrieval	Information sciences
Data wrangling	Information sciences and health sciences
Analytical modeling and machine learning	Information sciences
Healthcare workflow and reproducibility	Health sciences and management sciences
Communications, data visualization	Interdisciplinary
Data governance and ethical considerations	Interdisciplinary

Data literacy is a critical skill in many healthcare areas and is a required skill in health informatics. In health data science, the traditional term ***data literacy*** has evolved to refer not just to data management skills but also to **data acumen**, the ability to make good judgments about the use of data to support evidence-based decisions (Keller et al. 2020). The availability of open-source data science tools and libraries has made data science accessible to professionals outside computer science, which also has expanded the need for data acumen beyond data management. Table 1.3 lists data science skills and health data acumen that are gathered from various sources, including the seminal report from the National Academy of Sciences, Engineering, and Medicine on undergraduate data science education (NASEM 2018).

Check Your Understanding 1.3

1. Which one of these statements is incorrect?
   a. Information is data plus context that gives meaning
   b. Knowledge is interrelated pieces of information that form patterns
   c. Data have no meaning
   d. Information is larger than data
2. The DIKW pyramid is important because:
   a. It illustrates the relationship between data, information, and knowledge
   b. The pyramid reflects the degree of importance and volume
   c. Wisdom is the most important
   d. It helps understand the transformation of data to knowledge and wisdom

3. What are the three areas of skills of health informatics?
   a. Health sciences
   b. Information sciences
   c. Management sciences
   d. Public health
4. Which one of these is not one of the data acumen areas?
   a. Data management
   b. Analytical modeling and machine learning
   c. Advanced mathematics
   d. Communications and data visualizations

Chapter 1 Review Exercises

Review the following concepts:

1. Relationships between health informatics and health data science
2. How volume of data and new analytical methods can transform health informatics
3. The semantic gap that makes health data challenging
4. The DIKW pyramid in health data
5. The new skills of data acumen

REFERENCES

Ackoff, R. L. 1989. From data to wisdom. ***Journal of Applied Systems Analysis*** 16(1):3–9.

American Recovery and Reinvestment Act of 2009. Public Law 111-115.

Fridsma, D. B. 2018. Data sciences and informatics: What's in a name? ***Journal of the American Medical Informatics Association*** 25(1):109.

Keller, S. A., S. S. Shipp, A. D. Schroeder, and G. Korkmaz. 2020. Doing data science: A framework and case study. ***Harvard Data Science Review*** 2(1).

Kent, J. 2018. Big data to see explosive growth, challenging healthcare organizations. ***Health IT Analytics***. https://healthitanalytics.com/news/big-data-to-see-explosive-growth-challenging-healthcare-organizations.

LeSueur, D. 2017. Reasons healthcare data is unique and difficult to measure. ***Health Catalyst***. https://www.healthcatalyst.com/wp-content/uploads/2021/05/5-Reasons-Healthcare-Data-Is-Unique-and-Difficult-to-Measure.pdf.

National Academies of Sciences, Engineering, and Medicine. 2018. ***Envisioning the Data Science Discipline: The Undergraduate Perspective: Interim Report.*** Washington, DC: National Academies Press.

Ohno-Machado, L. 2013. Data science and informatics: When it comes to biomedical data, is there a real distinction? ***Journal of the American Medical Informatics Association*** 20(6):1009.

Payne, P. R. O., E. V. Bernstam, and J. B. Starren. 2018. Biomedical informatics meets data science: Current state and future directions for interaction. ***JAMIA Open*** 1(2):136–141. doi:10.1093/jamiaopen/ooy032.

Sardar, Z. 2020. The smog of ignorance: Knowledge and wisdom in postnormal times. ***Futures*** 120:102554.

Washington, V., K. DeSalvo, F. Mostashari, and D. Blumenthal. 2017. The HITECH era and the path forward. ***New England Journal of Medicine*** 377:904–906.

Chapter

2

Electronic Health Records and Clinical Information Systems

By Dilhari DeAlmeida, PhD, RHIA, and Suzanne Paone, DHA, MBA, RHIA

Learning Objectives

- Understand the evolution of the development of the electronic health record
- Analyze the issues surrounding the deployment and implementation of the electronic health record
- Contrast the differences in electronic health record systems
- Review the importance of the emerging digital health market
- Articulate the effects of regulations on the digital health market
- Summarize the advantages and disadvantages of the electronic health record
- Evaluate the current status and documented outcomes of electric health record utilization
- Articulate the role of the consumer health record

KEY TERMS

Ancillary systems
Behavioural healthcare
Certified Electronic Health Record Technology (CEHRT)
Clinical decision support
Clinical decision support system (CDSS)
Clinical documentation
Clinical information systems platforms (CIS)
Computerized provider (physician) order entry (CPOE)
Electronic health record (EHR)
Electronic medical record (EMR)
Electronic medication administration record (eMAR)

Emergency department information system (EDIS)
Episode of care
E-prescribing
Evidence-based medicine
Health information exchange (HIE)
Health Informatics (HI)
Health Information Technology for Economic and Clinical Health (HITECH) Act
Healthcare Information and Management Systems Society (HIMSS)
Integrated mobile devices
Integration
Interface
Interoperability
Internet computing
Laboratory information system (LIS)
Mainframe computing
Meaningful Use
Operating room information system (ORIS)
Personal health record (PHR)
Picture archiving and communication system (PACS)
Relational database systems
Technology platforms
Tethered PHR

Informatics involves informed decision-making on the part of critical stakeholders to improve the delivery of health-related services and the level of customer service provided to individuals in the role of patient in settings such as hospitals and clinics and in the role of consumer. Health provider organizations of all types, including health systems, hospitals, physician offices, clinics, long-term care facilities, rehabilitation facilities, retail care entities, and behavioural health providers, rely on **technology platforms**. These platforms generate reliable data that become knowledge-based information during the processes of delivering care. **Health informatics, HI**, is defined as using these platforms, data and clinical processes to address organizational needs. Clinical business processes such as documenting **episodes of care**, ordering and taking actions on clinical testing such as laboratory and radiologic exams, and administering safe medications are all enhanced by using technology platforms.

These platforms also add value by automating critical administrative processes necessary for the delivery of quality care such as insurance verification, resource-based scheduling, delivering telemedicine services, and the assimilation of patient and consumer-generated information. In the health field, new innovations such as medical devices that monitor health conditions and individuals' wellness allows for these platforms to expand to include patient generated information.

Technology platforms including software, servers, secure Internet services databases, and devices that administer clinical care and support clinical organizations' administrative processes are generally referred to as **clinical information systems platforms (CIS)**. Examples of CIS software systems include applications that support the laboratory information system (LIS), radiologic services (radiology information system, or RIS), and other clinical departments that use technology platforms in the delivery of care, such as cardiology and gastroenterology. CIS platforms include user interfaces for providers of care and technologists to use in the delivery of procedures and testing and share data with

medical equipment, such as testing analyzers in the laboratory, MRI machines, and nuclear cameras in cardiology. The platforms that integrate the data and functions of specialized CIS solutions in hospitals, ambulatory environments, long-term care environments, and consumers' homes are generally referred to as **electronic health record (EHR)** platforms. EHR solutions in health systems at hospitals also often include specialized models to support clinical administrative and supply chain functions in complex clinical areas, such as emergency rooms, operating rooms, perioperative environments, and cancer care centers.

EHR platforms involve comprehensive functions such as the ability for clinical providers to document care and interact with business intelligence at the point of medication orders and clinical testing. They also allow persons who use technology platforms, known as end users, in areas managing the revenue cycle such as insurance verification, admissions, and scheduling to facilitate organizational workflows, which optimizes the delivery of care. Many of these software platforms in health system, hospital, or independent physician office settings include modules to allow secure interactions with patients and consumers. Modules within the EHR that allow the secure transaction of data, services such as visits and scheduling, prescription management, and home device monitoring directly with the consumer are referred to as the **personal health record (PHR)**. Unlike software that can be downloaded from the Internet to track personal health, PHR software that is engineered into the database of an EHR is referred to as a **tethered PHR.**

The functions that define the EHR in the United States as well as the requirements providers of care must meet when using EHR platforms are regulated by federal requirements, administered by the Centers for Medicare and Medicaid Services (CMS), and tied to Medicare payment compliance. Likewise, commercial insurance companies that pay for health services often follow federal Medicare requirements regarding the use of the EHR to support value-based payment programs for their members. Additionally, many other functions within EHR and CIS platforms are regulated by mandated and voluntary accreditation programs, such as programs administered in the laboratory. For example, the College of American Pathologists, which regulates clinical laboratories, requires laboratory test data to be stored for a specified period of time. There are also federal compliance requirements and business requirements in regions of the US that mandate the sharing of data between different HR platforms. This concept of data sharing can be technical in nature, using programming interfaces, or manual by way of a consumer downloading a subset of their PHR. The automated or manual sharing of data between disparate EHR or PHR platforms is called **interoperability**. The major aim of EHR and CIS platform use and data sharing is the appropriate clinical care of consumers and improved efficiencies gained using these platforms in sponsoring organizations.

History

A historical perspective of the evolution of these platforms provides a valuable framework for excellence in informatics practice today. The history of EHR and CIS solutions involved early automation, primarily in hospitals and physicians' offices in the US. The evolution of these platforms was influenced by three external factors: the development of distributed computing and **internet computing**, societal concerns about the quality and cost of health care, and federal regulation developed and administered by CMS. Following early hospital and physician adoption, other settings of care such as long-term care, home care, retail care settings, and specialty care providers such as chiropractors, dental offices, and eye-care professionals transformed both clinical and administrative business practices using CIS and EHR systems. Since 2005, we have seen

a proliferation of internet-based delivery of clinical systems functionality such as personal health records, telemedicine, and **integrated mobile devices** that a patient might use to access a clinical health record.

As outlined in figure 2.1, a series of events reflecting changes in computing, attitudes toward quality medicine, and regulation shaped the development of CIS and EHR platforms. In the mid- to late 1980s, physicians' offices had practice management software for billing applications that automated claims to Medicare. Likewise, hospitals used **mainframe computing** to organize basic documentation of admissions, episodes of care, and testing to submit claims. Some companies offered mainframe-based nursing and unit clerk documentation systems in hospitals that were used to place orders for laboratory and radiology tests. These systems represented early order management and documentation functions.

Beginning in the 1990s, there was a proliferation of **relational database systems** and personal computing usage in many US industries, including healthcare. These technologies precipitated the development of advanced **ancillary systems** in the industry automating departments in hospitals, such as the laboratory, radiology, and pharmacy. The use of early personal computers in hospitals and physicians' offices also enabled vendors who worked with images to develop PC-based viewing technologies for **picture archiving and communication systems (PACS)**. This allowed the ubiquitous viewing of radiology and other images, such as EKGs outside of the radiology department, and significantly reduced operational challenges related to the printing and manipulation of expensive film-based images.

In the late 1990s and early 2000s, vendors emerged in the US software market offering not only ancillary systems, but also common relational database models for patient-centered electronic records. These early **electronic medical record (EMR)** platforms are called electronic medical records (EMR) because they focused on computing within hospital and physician office settings. They allowed hospitals to interface multiple ancillary systems into one patient-centered relational database with associated modules and applications that allowed providers to place orders and document care. Early template-driven documentation initiatives were sought after by hospital administrators as a method to improve turnaround time related to clinical documentation and reduce significant telephone-based transcription costs. This was also the period when specialty EMR applications developed in the market, such as applications that automated workflow and documentation in the emergency department **(emergency department information system, or EDIS)** and early operating room automation systems **(operating room information system, or ORIS)**. Later, these platforms expanded to include functions outside of hospitals and physician offices such as the ability to share data across different organizations and with patients. At this stage the systems began to be called EHR or electronic health record platforms.

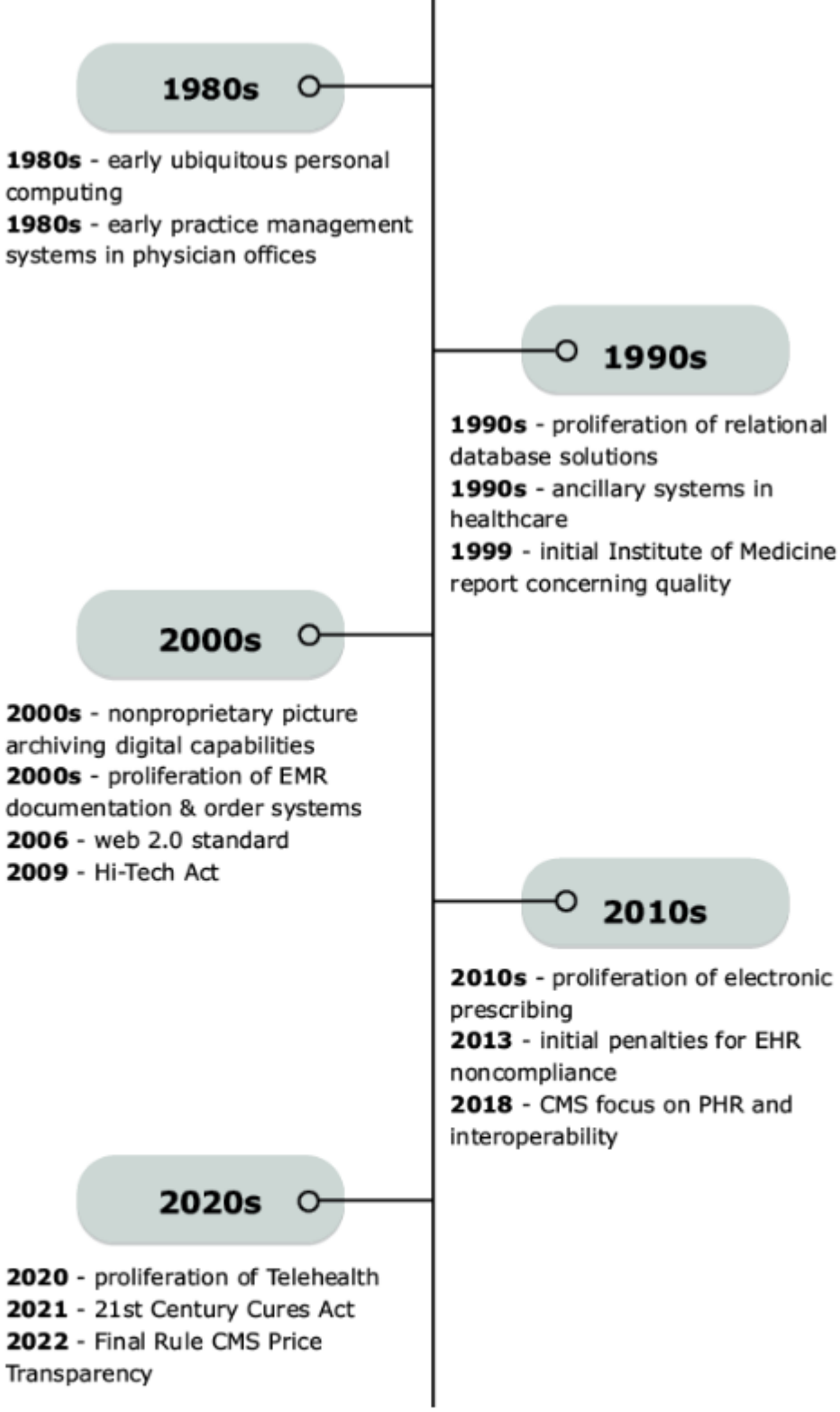

1980s
1980s - early ubiquitous personal computing
1980s - early practice management systems in physician offices
1990s
1990s - proliferation of relational database solutions
1990s - ancillary systems in healthcare
1999 - initial Institute of Medicine report concerning quality
2000s
2000s - nonproprietary picture archiving digital capabilities
2000s - proliferation of EMR documentation & order systems
2006 - web 2.0 standard
2009 - Hi-Tech Act
2010s
2010s - proliferation of electronic prescribing
2013 - initial penalties for EHR noncompliance
2018 - CMS focus on PHR and interoperability
2020s
2020 - proliferation of Telehealth
2021 - 21st Century Cures Act
2022 - Final Rule CMS Price Transparency

Figure 2.1. History of EHR

▶ Expand/collapse Extended Description

Source: Goodman, Kenneth W; Adams, Samantha, AMIA's Code of Professional and Ethical Conduct, *Journal of the American Medical Informatics Association*, 2013, 20(1), by permission of Oxford University Press.

Increases in insurance claims for testing processed by Medicare, the beginnings of price transparency, and six-sigma-based quality work in other industries ignited focus on quality and cost issues in the US healthcare system. The Institute of Medicine (IOM), now the National Academy of Medicine, organized working committees of medical professionals, private-sector companies, and academics to study and formulate recommendations regarding safety and quality in healthcare. In tandem, CEOs of sizable US-based corporations called for accountability in quality and cost controls related to escalating healthcare costs for privately insured employees. Six published aims of early IOM reports called for improved efficiency, safety, communication, and patient-centered care goals, implying the need for advanced digital adoption in the healthcare industry (Wolfe 2001). Escalating cost and quality issues in the US healthcare system called for industry accountability due to public concern. This led to federal legislation focusing on the use of information technology platforms in the healthcare industry.

A series of federally oriented legislative actions followed in 2009, beginning with the passing of the **Health Information Technology for Economic and Clinical Health (HITECH)Act**. This legislation empowered the Department of Health and Human Services (HHS) at the federal level to administer a program called **Meaningful Use (MU).** This program created tax-funded incentives for eligible providers of hospitals, physicians' offices, and clinics to adopt EHR platforms according to a series of phased published standards (HealthIT.gov n.d.). To receive financial incentives and payment, select providers and hospitals had to comply with phases of this program, which included incentives for critical safety-oriented CIS and EHR functions such as electronically legible orders, electronic prescribing, and legible and complete visit documentation. In early phases, providers received incentives for demonstrating secure networks and early adoption of critical EHR functions, such as order management and inclusion of social history-based issues in digital records. A good example of this is a patient's smoking status. As early as 2013, select providers received less revenue for noncompliance in the early phases of the MU program. This caused physicians' offices and hospitals to miss revenue goals that were planned within their budgets.

Later phases of the MU program required demonstrated use of messaging in a tethered PHR along with requirements for sharing data between disparate EHR platforms. Beginning in 2018, the focus of the program shifted to data sharing or industry-wide interoperability. Focus on interoperability continued with the implementation of the 21st Century Cures Act in 2021 (ONC 2021). This legislation not only promotes interoperability, but also creates significant financial penalties for individuals, provider organizations, and private-sector companies to participate in activities related to information blocking. For example, if a technology software company overprices services needed to share data for a hospital, this could be considered information blocking. Therefore, all stakeholders in the realm of selling, implementing, and consuming EHR systems, data, and services have an incentive to build platforms with open data sharing and cost-effective models allowing minimum data sets to flow across settings of care.

The MU program not only created financial incentives for EHR use and support for rural and underserved providers in system selection and implementation, but it also created a bureaucracy to certify EHR platform vendor software systems for a period. Today, the formal vendor certification program is no longer in place; however, significant industry consolidation and standardization of EHR software platforms occurred because of this legislation. Continued persistent administration of the Meaningful Use program beginning in 2009 through 2018 has resulted in consistent EHR platform adoption in the hospital, clinic, and physicians' office settings more than 90 percent of sites by 2018 (Mandeville et al. 2020). As of 2022, many hospitals and physicians' practices have implemented second- or third-generation EHR platforms as studies are beginning to

demonstrate an association between MU compliance proficiency and use of specific vendor products (Holmgren et al. 2018).

Many studies of CIS and EHR systems discussed the safety and operational efficiencies in the last 15 years. Even with all the benefits of EHR platforms and the use of a quality EHR system as a requirement of doing business, there continue to be challenges in the industry related to platform adoption and effective use of these tools. Qualitative academic studies and surveys performed by professional associations demonstrate a spectrum of ambivalence regarding the use of EHR systems with clinicians, a recognition that systems aid in the reduction of medical error, and a perception of the use of such systems for compliance purposes as time-consuming in the operational workflow of clinicians (Upadhyay and Hu 2022). US-based EHR companies leading the consolidated vendor market work on product development improvements to improve ease of use of EHR systems and reduce the time-consuming burden of documentation compliance rules represented in the use of the EHR software, including the integration of functions, such as voice recognition software.

The twenty-first-century expectations of EHR platforms include confidence that the utilization of these platforms and the data produced by these systems will assist in industry transformational efforts relating to improving consistent quality and reducing costs of healthcare. There is focused attention on issues that have been identified as gaps in the health system related to communication and coordination. Studies are finding associations between increased collaboration of clinical and administrative staff supported by EHRs. This includes the use of functions such as messaging, data sharing, and mutual awareness of the need for shared information and portable access EHR functions for specialists and nonspecialists (Vos et al 2020). As federal and private insurance entities shift payment models from fee-for-service to performance-based payments, organizations rely on EHR investments to ensure consistent quality delivery and manageable cost of services. Recent studies show that the implementation and effectiveness of EHRs as defined by MU is difficult to correlate with hospital value-based purchasing (HVBP) measures, which indicates that further studies are needed. In one article medication reconciliation and use of patient educational resources in EHR platforms showed a positive association with HVPB measures (Murphy et al. 2020). EHR and CIS platforms continue to play a critical role in the efficient and compliant operation of provider services, including supporting recent federal efforts requiring hospital price transparency, hospital reliance on data generated to develop service-based pricing, and meeting the Hospital Transparency Rule 2022 (CMS 2021).

Check Your Understanding 2.1

Answer the following questions.

1. Which of the following phrases most accurately describes the intention of the federal government's Meaningful Use (MU) program?
   a. A program providing support for rural doctors to receive dollars in their budgets for demonstrating effective use of the EHR
   b. A program providing support for select doctors and hospitals to receive dollars in their budgets for demonstrating effective use of the EHR
   c. A program designed to penalize physicians who do not purchase specific technology platforms
   d. A program designed to provide dollars in hospital budgets for demonstrating the effective use of administrative systems to reduce cost

2. Which is an example of legislation focused on utilizing HI to provide data to a hospital CEO supporting the cost of care?
   a. The Hospital Price Transparency Rule
   b. The Meaningful Use Program
   c. The HITECH Act
   d. Hospital value-based purchasing systems
3. Identify the HI function that was required of providers in the later (third phase) of the Meaningful Use program.
   a. Placing orders for clinical testing
   b. Accurately identifying a patient record
   c. Sending a finite number of messages to patients using the PHR
   d. Using the EDIS system to treat patients in the emergency department
4. The expectations for twenty-first-century EHR platforms are meant to improve and include the following:
   a. Improving consistent quality and driving down costs of healthcare systems
   b. Using technology to provide increased tests and services for patients
   c. Improving quality of care by increasing the costs that are rendered to consumers
   d. Studies have found decreased collaboration and communication among providers
5. As you review and discuss this section and figure 2.1 on the history of EHR, what are three important policies or pieces of legislation that are the most significant to you and why? Match the term with the appropriate description.
   a. Laboratory information system (LIS)
   b. Voice recognition
   c. 21st Century Cures Act
   d. HITECH Act
   e. Electronic prescribing
6. ___Technology developed to improve EHR usability for providers
7. ___Legislation creating financial incentives for EHR adoption
8. ___An example of an ancillary system
9. ___An example of an EHR function
10. ___Legislation that prohibits information blocking

Types and Components within an EHR

The EHR is not just software or a single computer system but is composed of multiple integrated technologies. These technologies might include, but are not limited to, software applications and platforms and various technologies. An integrated system is one that has been designed to bring together multiple information systems and allows them to communicate in a timely and effective manner and work together as one system. There might be variations among systems regarding which ancillary services are included. There also could be differences in the ways the interfaces provide access to the ancillary systems, but all EHRs are designed to create an electronic record to bring together information from disparate systems that are in compliance with standards for the purpose of supporting quality of care, administrative functions, and research. The federal government validates technology software vendor products and determines if they are appropriate for use. This determines whether healthcare providers receive reimbursement for medical services from the federal Medicare program. This process is referred to as certifying information technology. To use the EHR in a meaningful way for calendar year 2022, eligible hospitals and critical access hospitals (CAHs) may use the existing 2015

Edition certification criteria, the 2015 Edition Cures Update criteria, or a combination of the two to meet the **Certified Electronic Health Record Technology (CEHRT)** definition, as finalized in the CY 2021 PFS final rule (85 FR 84818 through 84828). ONC's 21st Century Cures Act Final Rule made several changes to the existing 2015 Edition Health IT Certification Criteria. The following changes constitute the 2015 Edition Cures Update:

- New technical certification criteria to advance interoperability and make it easier for patients to access their own electronic health information on their smartphones
- New privacy and security certification criteria
- Revised the standards referenced by several existing 2015 Edition certification criteria, including United States Core Data for Interoperability updates
- Removed and time-limited several 2015 Edition certification criteria

An EHR system is quite complex and supports more than the data needed for MU purposes. The systems are also used for capturing diagnostic and treatment information such as medications, lab, x-ray, and other clinical information in addition to the administrative uses.

The element that is identified as administrative in nature includes those data that accurately identify the patient, such as the name and other demographics, and the registration, admissions, discharge, and transfer information, including such things as chief complaint and patient disposition. A vital administrative consideration is the unique patient identifier, which is at the core of an EHR because it is the means by which all information from the various components is linked to the patient. For confidentiality and security purposes the patient identifier is unique to the organization and not known outside the organization.

In the majority of EHR systems, the laboratory systems remain as stand-alone systems that interface with the EHR and with numerous laboratory instruments that process the lab specimens. A **laboratory information system (LIS)** is a system for the integration of laboratory information related to the processing and storage of medical tests (physician orders for the tests and reports of results) that interfaces with the EHR to provide this information for inclusion in individual patient records. **Integration** involves ensuring that all aspects of an information system communicate appropriately and function as a uniform system (AHIMA 2017). Other LIS functionality includes scheduling, billing, and other information specific to the needs of the lab but not made a part of the EHR nor available to those with EHR access privileges.

Radiology information systems (RIS) bring together patient radiology data with images. As seen with laboratory systems, RIS systems often include patient tracking, scheduling, and reporting of results. Many systems are used with PACS. PACS provide storage of electronic images and reports. Although many hospitals have some sort of electronic system to support radiology, it does not mean that the system is necessarily integrated with the EHR.

Clinical documentation is "any manual or electronic notation (or recording) made by a physician, or other healthcare clinician, that is associated with a patient's medical condition and treatment" (AHIMA 2017). Structured or discrete data that are entered using a controlled vocabulary rather than narrative are referred to as structured data entry (AHIMA 2017). This would include patient assessments and clinical reports such as flow sheets and medication administration records. Conversely, clinical documentation entered into the patient record as text is narrative data and is not as easily automated due to the unstructured nature of the data. Examples of unstructured, narrative clinical information would include notes written by the physicians and other practitioners who treat the

patient, dictated and transcribed reports, and legal forms such as consents and advanced directives. Information from medical devices may also be integrated into the information flow such as vital sign readings or intravenous pump dosages and flow rates that can be sent directly to the patient record from the device where the data originated.

page 25

Table 2.1. Primary and secondary uses of an EHR system

Primary uses	Secondary uses
Patient care delivery	Education
Patient care management	Regulation
Patient care support processes	Research
Financial and other administrative processes	Public health policy and homeland security
Patient self-management	Policy support

Source: IOM 2003 (adapted from IOM 1997).

In 2003, a committee of the IOM identified potential primary and secondary uses of an EHR system as outlined in table 2.1. Primary uses are the direct recording of care on individual patients and the secondary uses involve aggregating patient data to improve the delivery of healthcare.

In addition to identifying the primary and secondary uses of information, the IOM outlined eight core functions that EHR systems should be capable of performing. These functions are designed to promote greater safety, quality, and efficiency in healthcare delivery. The eight functionalities with a brief description are:

1. Health information and data—information for care providers to make clinical decisions
2. Results management—management strategy using feedback for development to achieve goals
3. Order entry/management—"can improve workflow processes by eliminating lost orders and ambiguities caused by illegible handwriting, generating related orders automatically, monitoring for duplicate orders, and reducing the time to fill orders" (Lepage et al. 1992; Mekhjian et al. 2002; Sittig and Stead 1994)
4. Decision support—system designed to assist in making clinical decisions such as prescribing drugs, detection of adverse events and disease outbreaks, and computer prompts for such areas as vaccinations, breast cancer and colorectal screening, and cardiovascular risk reduction
5. Electronic communication/connectivity—communication among healthcare team members and other care partners (for example, laboratory, radiology, pharmacy) and with patients
6. Patient support—patient access to their health record and educational materials to engage the patient more in their healthcare and to participate in the clinical control of chronic illnesses

page 26

7. Administrative processes and reporting—to increase the efficiency of healthcare organizations and provide better, timelier service to patients with electronic scheduling, billing, and inventory
8. Reporting and population health management—reduce the significant data collection burden at the provider level with the associated costs and increase the accuracy of the

data reported with a standardized terminology and in a machine-readable format (IOM 2003)

See Table 2.2 for an example of the different clinical functions that might be present in an EHR.

Table 2.2. Examples of clinical functions in an EHR

Health information and data Patient demographic information Patient problem list Patient medication lists Clinical notes Minimum data set Notes including medical history and follow-up notes	Electronic communications and connectivity Electronic health information exchange (eHIE) Access to shared patient histories Continuity of Care Document (CCD)
Results management Viewing lab results Viewing imaging results Electronic images are returned	Patient support Patient portal to her (PHR) Email communication with clinicians
Order entry management Computerized orders for prescription Computerized orders for labs Computerized orders for radiology Orders sent electronically for prescriptions Orders sent electronically for labs Orders sent electronically for radiology	Administrative processes Scheduling/appointments Billing Inventory Medical coding
Decision support Warnings of drug interactions or contraindications are returned Out-of-range lab levels are highlighted Reminders for guideline-based interventions and screenings Access to online clinical guidelines	Reporting and population health management Disease reports Disease registries Social History tracking Quality measured and improvement reports Patient safety Immunization information exchange

Source: IOM 2003.

Check Your Understanding 2.2

Answer the following questions.

1. Which item is not a key function of the EHR?
   a. Clinical decision support in a doctor's office
   b. Hospital pharmacy orders
   c. Patients downloading their archived medication record
   d. Hospital payroll

2. Which is not an administrative process typically supported by the EHR in a hospital setting?
   a. Billing private insurance companies for inpatient care services
   b. Scheduling an MRI imaging test for a patient while in the hospital
   c. Staffing residents on the correct clinical rotations
   d. Maintaining the clinical drug inventories on inpatient units
3. Which is an example of clinical decision support?
   a. The EHR highlighting an out-of-range potassium result for a prescribing physician
   b. A nurse charting a patient's vital signs
   c. An order being sent to the pharmacy
   d. A training physician resident using the EHR to view a patient's latest white blood cell count
4. The ultimate goal of meaningful use is to:
   a. Engage patients
   b. Improve health outcomes and health of populations
   c. Provide data for aggregate data analysis
   d. Provide monetary support to physicians for obtaining EHR systems
5. Which of the following is a primary use of an EHR system?
   a. Patient care delivery
   b. Research
   c. Patient self-management
   d. Financial processes

Inpatient EHR

As described earlier in this chapter, inpatient EHRs have evolved over time, beginning with business- and financial-based hospital information systems, followed by clinical information systems, and evolving to EHRs. Inpatient EHRs are generally comprised of integrated, interoperable multiple applications. In contrast, most ambulatory EHRs have been developed as complete systems that include practice management and clinical applications. An inpatient EHR's components could include financial and administrative applications; clinical systems; **Computerized Physician Order Entry (CPOE)**, a process where a clinician enters and sends care instructions via electronic system; **electronic medication administration records (eMARs)** (use of technology and bar coding to track medications from when ordered to when given to the patient); clinical data repositories; **clinical decision support**; document imaging; and PACS.

Inpatient EHRs certified for MU purposes may be classified as complete or modular. Complete systems are those that meet all of the MU criteria tested for CEHRT certification. Certified modular systems meet one or more of the MU criteria so would require multiple modular systems working together to attest compliance with Stage 1. According to the ONC, data gathered from 2019 and 2021 show the adoption rates for EHR in various different healthcare settings: acute care, psychiatric, specialty, rehabilitation, children's, and acute long-term care (ONC 2022). Eighty-six percent of non-federal general acute care hospitals had adopted a 2015 Edition certified EHR. In contrast, only 40 percent of rehabilitation hospitals and 23 percent of specialty hospitals had adopted a 2015 Edition certified EHR. Extensive resources and effort have been put into implementing hospital-setting EHR systems. There are administrative and financial systems in the EHR system that include R-ADT for registration and admission discharge

and transfer; MPI, the master patient index; and PFS, practitioner/physician fee schedule. The ancillary/clinical and specialty clinical applications provide data when they are used for an individual patient. For example, if a patient had an x-ray, the radiology system would transmit the report to become a part of the EHR. For patients who were in the ICU, that system would provide reports specific to the time there. The "smart" peripherals are the devices that can submit reports directly from the device to the EHR, such as EKG equipment monitoring a patient's heart function transmits reports that become a part of the record. With all of this patient information in the patient record, it will then be used to carry out the core clinical applications to treat and document the care given to the patient. Examples of this are the point-of-care recommendations (POC), CPOE for medications, diagnostic tests and other things that must be ordered for the patient, and clinical decision support information for provider to use for reference. The other parts of the system are those required for making the system function.

The **Health Information and Management Systems Society (HIMSS)** EMR Adoption Model (EMRAM[SM]) provides a framework for describing the components of an EHR and measuring and reporting the stage of adoption by healthcare organizations (see table 2.3). Table 2.3 depicts the HIMSS adoption model that can serve as a benchmark for others in their pursuit of implementing technology (HIMSS 2016). EHR systems have increased in numbers and the level of use of the information available with the use of these systems. Funding initiatives, reimbursement considerations, and regulations are the reasons for the increased deployment of EHRs. The extent to which systems have adopted the potential system components varies. Using the HIMSS EMR Adoption Model as the basis of measuring the level of adoption, this variation is quite evident. the case study that follows shows the adoption level of EMRs based on the seven stages identified by HIMSS.

Table 2.3. HIMSS analytics adoption model

Stage 0	The organization has not installed all of the three key ancillary department systems (laboratory, pharmacy, and radiology).
Stage 1	All three major key ancillary clinical systems are installed.
Stage 2	Major ancillary clinical systems feed data to a clinical data repository (CDR) that provides physician access for reviewing all orders and results. The CDR contains a controlled medical vocabulary, and the clinical decision support/rules engine (CDS) for rudimentary conflict checking. Information from document imaging systems may be linked to the CDR at this stage. The hospital may be health information exchange (HIE) capable at this stage and can share whatever information it has in the CDR with other patient care stakeholders.
Stage 3	Nursing/clinical documentation (for example, vital signs, flow sheets, nursing notes, eMAR) is required and is implemented and integrated with the CDR for at least one inpatient service in the hospital; care plan charting is scored with extra points. The Electronic Medication Administration Record application (eMAR) is implemented. The first level of clinical decision support is implemented to conduct error checking with order entry (that is, drug–drug, drug–food, drug–lab conflict checking normally found in the pharmacy information system). Medical image access from PACS is available for access by physicians outside the radiology department via the organization's intranet.
Stage 4	CPOE for use by any clinician licensed to create orders is added to the nursing and CDR environment along with the second level of clinical decision support capabilities related to evidence-based medicine protocols. If one inpatient service area has implemented CPOE with physicians entering orders and completed the previous stages, then this stage has been achieved.

Stage 5	A full complement of radiology PACS systems provides medical images to physicians via an intranet and displaces all film-based images. Cardiology PACS and document imaging are scored with extra points.
Stage 6	Full physician documentation with structured templates and discrete data is implemented for at least one inpatient care service area for progress notes, consult notes, discharge summaries or problem list and diagnosis list maintenance. Level three of clinical decision support provides guidance for all clinician activities related to protocols and outcomes in the form of variance and compliance alerts. The closed loop medication administration with bar-coded unit dose medications environment is fully implemented. The eMAR and bar coding or other auto identification technology, such as radio frequency identification (RFID), are implemented and integrated with CPOE and pharmacy to maximize point-of-care patient safety processes for medication administration. The "five rights" of medication administration are verified at the bedside with scanning of the bar code on the unit dose medication and the patient ID.
Stage 7	The hospital no longer uses paper charts to deliver and manage patient care and has a mixture of discrete data, document images, and medical images within its EMR environment. Data warehousing is being used to analyze patterns of clinical data to improve quality of care, patient safety, and care delivery efficiency. Clinical information can be readily shared via standardized electronic transactions (i.e., CCD) with all entities that are authorized to treat the patient, or a health information exchange (i.e., other nonassociated hospitals, ambulatory clinics, subacute environments, employers, payers, and patients in a data sharing environment). The hospital demonstrates summary data continuity for all hospital services (e.g., inpatient, outpatient, ED, and with any owned or managed ambulatory clinics). Blood products and human milk are included in the closed-loop medication administration process.

Source: Adapted from HIMSS 2013.

Case Study: Stage 7 EHR Adoption

General Health System (GHS) is comprised of three hospitals, 100 physician practices, a long-term care rehabilitation facility, and an ambulatory medical neighborhood convenience facility in the Midwest US. GHS implemented a comprehensive, enterprise EHR platform across hospitals and physicians and medical neighborhoods and a best-of-breed CIS system that specializes in rehabilitation in their rehabilitation facility. There is a two-way digital insurance, registration scheduling orders, and results system interface between the rehabilitation CIS system and the enterprise EHR. All documentation, medication administration, and orders are managed through the enterprise EHR. GHS is part of the original electronic prescribing network, sending prescriptions to retail pharmacies and checking formularies for medication pricing. The laboratory radiology and pharmacy modules as well as the operating room and emergency room models are all implemented and fully integrated in the enterprise EHR.

GHS implemented the tethered PHR kiosks in physician practices for standard check-in and checkout processes. Difficult-to-fill office staff positions have been reduced by 15 percent and standard customer flow through offices shows a decrease in wait time of 15 percent. GHS implemented standard glucometer scales and blood pressure monitors for their patients with diabetes, congestive heart failure, and hypertension. A technology consulting team is building an interface that will allow customers

of GHS to upload data from their iOS or Android health monitoring devices through the tethered PHR into the EHR on a nightly basis. The CMS Standard fields for clinical and administrative interoperability requirements are uploaded to the regional health information exchange (HIE). Likewise, nightly file uploads are exchanged with the state vaccination and public health system and three major retail pharmacy chains.

Consumers of GHS ambulatory or medical neighborhoods enjoy fully integrated telemedicine visits for episodic illness as well as telemedicine visits for post-surgery follow-up and hypertension management. Last year, GHS implemented online bill payment through their tethered PHR. Forty-five percent of co-pays and balances due are paid online, decreasing expenses associated with paper statements by $4 million per year and days in accounts receivable by two days, resulting in a 32 percent improvement in cash flow for the organization.

Ambulatory EHR

Ambulatory and inpatient EHRs have evolved differently over time based on the way software vendor products were sold in the marketplace. The specific components or functions also vary between ambulatory and inpatient EHRs. Ambulatory care is provided in a variety of locations by individual or groups of healthcare providers. Settings for ambulatory care may include physicians' offices, outpatient clinics, ambulatory surgery centers, urgent care clinics, retail clinics, and stand-alone laboratory and imaging facilities. In the ambulatory setting, the patient generally moves themselves from setting to setting when receiving services from multiple providers. Inpatient care is generally provided in the hospital setting with many levels of providers and services concentrated within one setting. EHRs in the ambulatory setting are similar to those seen in the inpatient environment. Major differences between these two kinds of EHRs are that the ambulatory record is longitudinal rather than episodic, has fewer practitioners documenting in the record, may not have in-house diagnostics (lab, x-ray, and such) and so requires a link to this information, and has more referrals made to outside practitioners than seen in the hospital. Ambulatory EHRs include:

- Problem lists
- Medication lists
- History and physical examinations
- Immunization records
- Patient progress and visit notes
- Clinical templates
- Electronic prescribing
- Medication management
- Health maintenance
- Disease management
- Outcomes reporting
- Clinical decision support
- Lab order entry
- Results reporting

Long-Term Care EHR

EHR systems in long-term care (LTC) facilities provide the same benefits as do those for acute care and will allow these facilities to participate in the exchange of data. Most LTC facilities have submitted the federally required Minimum Data Set (MDS) electronically for some time (AHIMA 2017). Other electronic functionality is limited in this environment primarily due to lack of financial and technical support. Future changes for LTC would most likely be processing changes rather than functionality changes. Specifically, processing changes in chart processing (assembly, auditing), moving and thinning of records and release of information (electronic exchange and documentation of disclosures), and reports and screen design could optimize processing time and output and provide more storage space and greater access to information. Practical LTC activities that would be enhanced with an EHR include care coordination with access to information that is already available, facilitating communication with all of the providers associated with the patient and for quality improvement to protect this high-risk group with prevention of additional conditions and earlier interventions when they do occur. Forces that are behind the desire for the EHR in LTC include quality of care, consumerism, and provision of care in a more efficient and effective way.

Behavioral healthcare can be provided in the ambulatory or inpatient settings. **Behavioral healthcare** is a "broad array of psychiatric services provided in acute, long-term, and ambulatory care settings; includes treatment of mental disorders, chemical dependency, mental retardation, and developmental disabilities, as well as cognitive rehabilitation services" (AHIMA 2017). Although the development and use of EHR for behavioral healthcare services has been slower to occur, a more recent report published by Fortune Business Insights states the behavioral healthcare market is predicted to grow from $77.62 billion in 2021 to $99.40 billion in 2028. Therefore, there is an increased need for using EHR for behavioral healthcare services such as delivering high-quality care, automating work processes, and assisting in streamlining communication between patients and clinicians, saving time and money. Healthcare staff could utilize the EHR for appointment reminders to patients through text, email, or phone call, and e-prescribing, to ensure that patients receive the accurate medication on time. **E-Prescribing** is a process by which medication orders and insurance information are digitally passed between physicians, pharmacies, and insurance companies. As mentioned before, these behavioral EHR systems could reside within an acute care or inpatent EHR if it is within a larger healthcare organization, or it could be a stand-alone EHR within an ambulatory setting. If an organization needs to use the EHR for behavioral health purposes, it would need to determine what works best for the clinicians who deliver the care and the patients who receive care. Another growing area of importance related to behavioral health is the evaluation of the social determinants of health. Social determinants such as poverty, which often results in frequent moves and changes in schools, limited health care, overcrowding, and unsafe environments, are acknowledged as potent risk factors for behavioral health disorders. Because of this, there is expected to be an even larger growth in the usage of the behavioral health EHR systems in the coming decades.

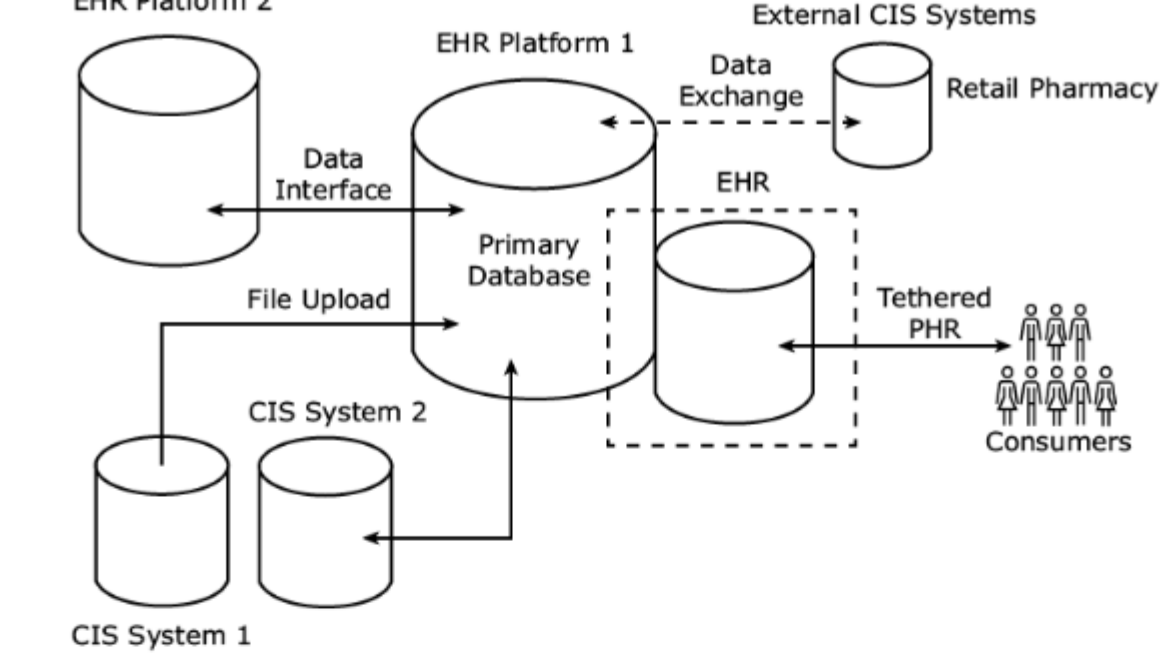
Personal Health Record

According to HealthIT.gov, there are two main types of personal health records (PHRs): standalone PHRs and tethered or connected PHRs. With a stand-alone PHR, consumers fill in information from their own records and the information is stored on their own computers or the internet. The consumer has the responsibility to accurately document their own healthcare needs and share that information with the provider. In some cases, a

stand-alone PHR can also accept data from external sources, including providers and laboratories. With a stand-alone PHR, consumers can add diet or exercise information to track progress over time. Consumers have the freedom to decide whether to share the information with providers, family members, or anyone else involved in their care (https://www.healthit.gov/faq/are-there-different-types-personal-health-records-phrs).
Tethered or connected PHRs are linked to a specific health care organization's EHR system or to a health plan's information system. With a tethered PHR, consumers can access their own records through a secure portal and view, for example, the trend of their lab results over the last year, their immunization history, or due dates for screenings (https://www.healthit.gov/PHR). An example of a tethered PHR is shown in figure 2.2. Often, a PHR that is tied to an EHR is called a patient portal. This concept is growing in the EHR/PHR space where healthcare organizations are giving access to consumers to manage and track certain aspects of their healthcare through a patient portal. Consumers can also use various health apps on their mobile devices. For example, the Medical ID for iPhones makes critical information available via the lock screen for use by first responders in an emergency. There are many advantages to using a PHR, however, there are some drawbacks to consider as well.

Advantages to using the tethered PHR for health systems, hospitals, and physicians' offices and other ambulatory sites fall into three specific categories: regulatory compliance with the interoperability phase of the CMS MU program, organizational expense management and efficiencies, and chronic disease management mandates and increased consumer engagement. Table 2.4 outlines some of the functions of tethered PHRs and organizational and customer benefits of using these software functions as well as areas of potential implementation concern.

In summary, this chapter introduced the reader to the advancements in the EHR over the years. Some of the key highlights and collaborations that have occurred between healthcare organizations, providers, and other key stakeholders in working within the EHR have proven to be promising. The different types of EHRs and their components are useful in identifying and treating consumers in a more effective manner. They have also given consumers the power to manage their own health records. The next chapter focuses on the strategies that are currently being used to connect different EHR systems.



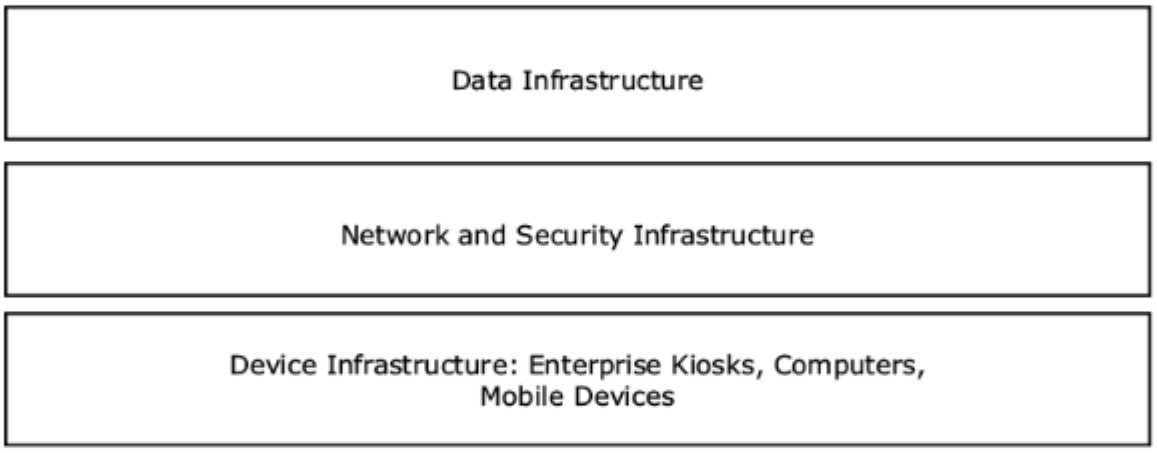

Data Infrastructure
Network and Security Infrastructure
Device Infrastructure: Enterprise Kiosks, Computers, Mobile Devices

Figure 2.2. EHR-tethered PHR portal

▶ Expand/collapse Extended Description

Table 2.4. Functions, benefits, and challenges to consider when using tethered PHR

Sample tethered PHR functions	Benefits	Potential challenges to consider
Automatic release of EHR test results	Satisfied customers Top reason people use tethered portals Reduced delays in treatment	Physician cultural considerations related to direct release of test results New policies for clinical turnaround time for results review

Co-negotiated medications, vaccine and problem lists between consumers and providers	Improved clinical safety Increased completeness of health record Providers and consumers agree on adding data to EHR from nonnative care sites (e.g., retail care, other hospitals) Increased consumer awareness/health literacy	Offices must adapt policies and workflows to allocate clinical decision makers to commit nonnative PHR data to the EHR record Offices need to educate and engage consumers about what they see on problem lists (e.g., codes such as obesity or depression)
PHR pathways and health maintenance reminders	Chronically ill consumers receive evidence-based care in a timely fashion Health organizations receive improved incentive payments from CMS and insurance companies via accountable care organizations and value-based payment contracts	Organizations must agree on building standard pathways in EHR based on accountable care and/or payment incentive requirements from CMS and/or commercial insurance companies
PHR integrated biometrics and health monitoring	Consumers engage providers in health management Consumers on chronic disease management care pathways such as diabetes or congestive heart failure submit to fewer hospital readmissions	Organizations must either adopt standardized devices and/or build interfaces from stand-alone devices through open system EHR platform toolkits
Nonvisit messaging to offices	Consumer convenience accessing data (e.g., vaccine records, verification for testing, increased access to clinicians) Competitive benefit for ambulatory sites competing with retail care and urgent care sites for consumer visits	Offices must adapt workflows to address administrative consumer needs and nonurgent medical inquiries
Billable tethered PHR telemedicine visits	Convenience for consumers Safety in visits executed with complete EHR data Tethered record data meets visits criteria for 99444 series billable CPT services	Offices need to adapt to new workflow and service models to deliver telemedicine in a timely fashion

PHR implementations of office tablet and/or kiosk software and online scheduling	Convenience for consumers visiting physical offices; check-in and checkout expedited Consumer satisfaction with online scheduling; second most frequently used tethered PHR function Offices save administrative costs of office staff	Administrative staff must work to reorganize workflows Clinicians must open certain schedules online
Tethered PHR implementation of bill payment online	Consumer convenience- level of service expected in other industries Significant cost reductions in processing paper statements; pays for tethered PHR licensing costs Significant reductions in days in accounts receivable improving cash flow and organizational financial health	Organizations must build interfaces to and from billing systems
Tethered PHR proxy functions	Consumers' caregivers manage administrative and clinical health of loved ones from anywhere Improved discharge and continuum of care planning for case managers	Health information management processes must be digitized to approve and manage proxy stakeholders Clinicians may need to educate family or other caregivers on content of the digital health record

Chapter 2 Review Exercises

Discussion

1. Research the landmark publication that came out in 1999 titled "To Err Is Human: Building a Safer Health System" that addressed the quality of the US healthcare system (you can find the article through Pubmed or similar sources online). What are four key factors that the report addressed in contributing to medical errors?
2. Learning about the history of the EHR and consulting figure 2.1, what do you consider the most significant milestone for each decade? Please make sure to provide justification for your answer. You can use additional resources and personal experience in answering the question.
3. As federal and private insurance entities shift payment models from fee-for-service to performance-based payments, organizations rely on EHR investments to ensure consistent quality delivery and manageable cost of services. Research an article that discusses the transition from fee-for-service to performance-based payments and discuss three things that conclusions from that article. You can discuss benefits/challenges or compare and contrast the items of discussion.

4. Explain the nature of the relationship between provider data and patient/client data in the tethered PHR environment.
5. Name one other setting of care where EHRs are used other than hospitals and physician offices.
6. Evaluate the healthy people 2030 initiative project at https://health.gov/healthypeople and share some of the and review some of the priority areas that they discuss, specifically discuss your understand of the social determinant of health and how you think it may reduce healthcare disparities.
7. Research on how electronic health records are starting to incorporate social determinants of health, behavioral health and disabilities into the electronic health record. Research at least 1-2 articles for each of these topics.
8. Discuss the different types of PHR systems and compare and contrast them to mobile healthcare apps that you may be using. What are your thoughts of each system as it relates to usability, storage of information etc.
9. Research the landmark publication "Crossing the quality chasm" that addressed the quality of the US healthcare system (you can find the article through PubMed or similar sources online). What are the key factors that the report addressed in contributing to improving quality?
10. Although the EHR has many advantages and areas to improve the quality of care, the user interface may not be optimal for clinicians. Often times, clinicans may find themselves moving through or clicking on many different alerts/screen/reminders etc. What are some suggestions of reducing clinical alert fatigue within the clinical decision support systems in an EHR?

11. Which is not an example of the benefit of a tethered PHR?
    a. Allowing health system customers to pay their co-pays online
    b. Providing a secure method for an approved daughter of an elderly patient to message the primary care physician's office concerning a new clinical problem
    c. Providing a secure method for a patient to place a refill for their prescription allergy medicine
    d. All of the above

REFERENCES

Adair, D. 2022. "Operational considerations for patient generated health data." https://journal.ahima.org/page/operational-considerations-for-patient-generated-health-data.

AHIMA. 2017. ***Pocket Glossary of Health Information Management and Technology,*** 5th ed. Chicago: AHIMA Press.

Amatayakul, M. K. 2013. Electronic Health Records: Conceptual Framework. Chapter 5 in ***Health Information Management: Concepts Principles and Practice***, 4th ed., edited by K. M. LaTour, S. Eichenwald Maki, and P. Oachs. Chicago: AHIMA.

Bush, G. 2004. "The 2004 State of the Union Address: Complete transcript of President Bush's speech to Congress and the nation." http://whitehouse.georgewbush.org/news/2004/012004-SOTU.asp.

Centers for Medicare and Medicaid Services (CMS). 2016. "Electronic health records incentive programs." https://www.cms.gov/Regulations-and-Guidance/Legislation/EH

RIncentivePrograms/index.html?redirect=/ehrincentiveprograms.
Centers for Medicare and Medicaid Services (CMS). 2021. "Hospital price transparency." https://www.cms.gov/hospital-price-transparency.
Department of Health and Human Services, Agency for Healthcare Research and Quality (AHRQ). 2006. "Barriers to HIT implementation." https://healthit.ahrq.gov/health-it-tools-and-resources/health-it-costs-and-benefits-database/barriers-hit-implementation.
Fleming, N. S., S. D. Culler, R. McCorkle, E. R. Becker, and D. J. Ballard. 2011. The financial and nonfinancial costs of implementing electronic health records in primary care practices. ***Health Affairs*** 30(3):481–489. http://content.healthaffairs.org/content/30/3/481.full.
Health Story Project. 2012. https://healthstoriesproject.com/.
HealthIT. 2012. "How much is this going to cost me?" https://www.healthit.gov/providers-professionals/faqs/how-much-going-cost-me.
HealthIT. 2013. "EHR incentives and certification." https://www.healthit.gov/providers-professionals/ehr-incentive-programs.
HealthIT. 2015. "Meaningful use definitions and objectives." https://www.healthit.gov/providers-professionals/meaningful-use-definition-objectives.
Hersh, B. 2015. "The conundrum of structured vs. unstructured data." http://informaticsprofessor.blogspot.com/2015/02/the-conundrum-of-structured-vs.html.
HIMSS. 2005. "Interoperability definition and background." http://www.himss.org/himss-interoperability-definition-and-background.
HIMSS. 2016. "U.S. EMR adoption model trends." https://app.himssanalytics.org/stagesGraph.asp.
HIMSS Analytics. 2015. "Arcare Stage 7 case study." https://www.himssanalytics.org/case-study/arcare-stage-7-case-study.
Institute of Medicine (IOM), Committee on Data Standards for Patient Safety. 2003. ***Key Capabilities of an Electronic Health Record System.*** Washington, DC: National Academy Press.
Institute of Medicine (IOM), Committee on Improving the Patient Record. 1997. ***The Computer-Based Patient Record: An Essential Technology for Health Care.*** Rev. ed. Washington, DC: National Academy Press.
Institute of Medicine (IOM), Committee on Quality of Health Care in America. 2001. ***Crossing the Quality Chasm: A New Health System for the 21st Century***. Washington, DC: National Academy Press.
Lepage, E. F., R. M. Gardner, R. M. Laub, and O. K. Golubjatnikov. 1992. Improving blood transfusion practice: Role of a computerized hospital information system. ***Transfusion*** 32(3):253–259.
Lippeveld, T., R. Sauerborn, and C. Bodart. 2012. Health Information Systems, Section 3, Module 7 in ***Health Systems Approach: A How-to Manual***. Arlington, VA: Management Sciences for Health.
McCullough, J. 2008. Adoption of hospital information systems. ***Health Economics*** 17:649–664.
Mekhjian, H. S., R. R. Kumar, L. Kuehn, T. D. Bentley, P. Teater, A. Thomas, B. Payne, and A. Ahmad. 2002. Immediate benefits realized following implementation of

physician order entry at an academic medical center. ***Journal of the American Medical Informatics Association*** 9(5):529–539.

Murphy, Z. R., J. Wang, and M. V. Boland. 2020. Association of electronic health record use above meaningful use thresholds with hospital quality and safety outcomes. ***JAMA Network Open*** 3(9):e2012529. https://doi.org/10.1001/jamanetworkopen.2020.12529.

National Academies. 2016. "About the IOM." http://www.nationalacademies.org/hmd/.

Office of the National Coordinator for Health Information Technology (ONC). 2014. "Report to Congress. Update on the adoption of health information solution technology and related efforts to facilitate the electronic use and exchange of health information."

Office of the National Coordinator for Health Information Technology (ONC). 2015. "Data Brief No. 28. Any, certified, and basic: Quantifying physician EHR adoption."

Office of the National Coordinator for Health Information Technology (ONC). 2016. "Hospital progress to meaningful use by size, type, and urban/rural location." Health IT Quick-Stat #5. http://dashboard.healthit.gov/quickstats/pages/FIG-Hospital-Progress-to-Meaningful-Use-by-size-practice-setting-area-type.php.

Office of the National Coordinator for Health Information Technology (ONC). 2021. 21st Century Cures Act Final Rule. https://www.healthit.gov/curesrule/.

Office of the National Coordinator for Health Information Technology (ONC). 2022. "Adoption of electronic health records by hospital service type 2019-2021, Health IT Quick Stat #60." https://www.healthit.gov/data/quickstats/adoption-electronic-health-records-hospital-service-type-2019-2021.

Office of the National Coordinator for Health Information Technology (ONC)/American Hospital Association (AHA). AHA Annual Survey Information Technology Supplement, 2019 & 2021. https://www.ahadata.com/system/files/media/file/2020/10/2019_AHAIT_Survey_1.pdf

Regenstrief Institute. 2016. http://www.regenstrief.org.

Sittig, D. F., and W. W. Stead. 1994. Computer-based physician order entry: The state of the art. ***Journal of American Medical Informatics Association*** 1(2):108–123.

Smithwick, J. 2015. "Unlocking the Value of Unstructured Patient Data." Becker's Health IT and CIO Review. http://www.beckershospitalreview.com/healthcare-information-technology/unlocking-the-value-of-unstructured-patient-data.html.

Tripathi, M. 2012. EHR evolution: Policy and legislation forces changing the EHR. ***Journal of AHIMA*** 83(10):24–29.

U.S. Department of Veterans Affairs. 2015. "E-Health." http://www.ehealth.va.gov/vista.asp.

Upadhyay, S., and H. Hu. 2022. A qualitative analysis of the impact of electronic health records (EHR) on healthcare quality and safety: Clinicians' lived experiences. ***Health Services Insights***.1–7. https://doi.org/10.1177/11786329211070722.

Vos, J. F. J., A. Boonstra, A. Kooistra, M. Seelen, and M. van Offenbeek. 2020. The influence of electronic health record use on collaboration among medical specialties. ***BMC Health Services Research*** 20(1):1–11. https://doi.org/10.1186/s12913-020-05542-6.

HealthIT.gov, (n.d.) ***What Is Meaningful Use.*** Retrieved from: https://www.healthit.gov/faq/what-meaningful-use.

Wolfe, A. 2001. Institute of Medicine report: Crossing the quality chasm: A new health care system for the 21st century. ***Policy, Politics & Nursing Practice*** 2(3). https://doi.org/10.1177/152715440100200312.
https://www.healthit.gov/faq/are-there-different-types-personal-health-records-phrs
https://www.healthit.gov/curesrule/final-rule-policy/2015-edition-cures-update
https://www.healthit.gov/faq/are-there-different-types-personal-health-records-phrs
Promoting Interoperability Programs (n.d.) Retrieved from: https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms

ADDITIONAL RESOURCES

AHIMA. 2014 Update: AHIMA Longitudinal Coordination of Care Practice Council E-HIM Strategy Team. ***Electronic Health Record Adoption in Long Term Care***. AHIMA Practice Brief.

Centers for Disease Control and Prevention (CDC). 2012. "National perceptions of EHR adoption: Barriers, impacts, and federal policies." http://www.cdc.gov/nchs/ppt/nchs2012/SS-03_JAMOOM.pdf.

Charles, D., M. Gabriel, and T. Searcy. 2015. Adoption of Electronic Health Record Systems Among U.S. Non-Federal Acute Care Hospitals: 2008–2014. ONC Data Brief, No. 23. Office of the National Coordinator for Health Information Technology, Washington, DC.

Department of Health and Human Services (HHS). 2012. "Press release: More doctors adopting EHRs to improve patient care and safety." https://content.govdelivery.com/accounts/USHHS/bulletins/61c93a.

Dick, R. S., and E. B. Steen. 1991. ***The Computer-Based Patient Record: An Essential Technology for Healthcare***. Washington, DC: National Academy Press.

HIMSS. 2004. Top line: President Bush calls for EHRs for most Americans in 10 years; creates new NHII coordinator. ***Health IT News***. http://www.healthcareitnews.com/news/president-bush-continues-ehr-push-sets-national-goals.

Jha, A., C. DesRoches, E. Campbell, K. Donelan, S. Rao, T. Ferris, A. Shields, S. Rosenbaum, and D. Blumenthal. 2009. Use of electronic health records in U.S. hospitals. ***The New England Journal of Medicine*** 360:1628–1638.

Sackett, D. 1996. Evidence based medicine: What it is and what it isn't. ***British Medical Journal*** 312:71–72. https://www.ncbi.nlm.nih.gov/pubmed/8555924.

Chapter

3

Healthcare Standards, Classification Systems, and Healthcare Information Exchange

By Dilhari DeAlmeida, PhD, RHIA, and Andi Saptono, PhD

Learning Objectives

- Identify and compare the major data sets, classification systems, clinical terminologies, and other standards utilized for secondary data use
- Interpret health information exchange concepts
- Identify the benefits of being able to exchange information across health entities
- Explain the benefits of interoperability for population and public health

KEY TERMS

Accountable Care Organization (ACO)
Aggregated HIE
Classification system
Cooperative HIO
Data segmentation
Federated HIE
Health information exchange (HIE)
Health information exchange organization (HIO)
Healthcare code sets
Healthcare Common Procedure Coding System (HCPCS)
International Classification of Diseases, 10th revision, Clinical Modification (ICD-10-CM)

International Classification of Diseases, 10th revision, Procedure Coding System (ICD-10-PCS)
Linearization
Logical Observation Identifiers Names and Codes (LOINC)

Mortality
No consent model
Opt-in model
Opt-in with restrictions model
Opt-out model
Opt-out with exceptions model
Population health
Private HIO
Public health
Public HIO
RxNorm
SNOMED CT
Standards
Terminology
Vocabulary

Healthcare data is becoming more and more prevalent among consumers as we start collecting, analyzing and disseminating it for different purposes. Public health reporting, reimbursement, and quality improvement are purposes for collecting healthcare data. However, there are many other areas that could use healthcare data for improving quality practices or provided better patient care or improving patient experiences. Such examples include provider certification or accreditation; disease management; determining best practices or clinical guidelines; detecting financial fraud; monitoring patient compliance; and identifying candidates for clinical trials or using electronic health record (EHR) data for clinical trials. In fact, one of the phrases often heard is "collect once, use many" to describe healthcare data that are collected for patient care purposes and then used to meet a myriad of additional requirements. Organizations describe "big data" to convey the message that the widespread implementation of health information technology (HIT) has resulted in extraordinarily large and detailed databases being made available for secondary data use.

Because healthcare is a very complex business, there are many concepts and ideas that must be accessed, combined, manipulated, and shared. Healthcare information systems use **standards** to describe these concepts and ideas in the form of vocabularies, terminologies, and classification systems. A **vocabulary** is a dictionary of terms; that is, words and phrases with their meanings. A **terminology** is a set of terms representing a system of concepts, usually around a specific domain such as healthcare. Finally, a **classification system** arranges or organizes like or related entities (Giannangelo 2015). This section explores code sets, classification systems, clinical terminologies, and other data set standards important to health informatics (HI) and the interoperability of health information. The process of sharing important health information is also discussed from the perspective of the exchange process itself, also known as **health information exchange**, or more popularly known as **HIE**, and the perspective of the entity that manages and ensures the success of the exchange, also known as the **health information exchange organization**, or **HIO**.

Healthcare Code Sets

Healthcare code sets are defined as "any set of codes used for encoding data elements, such as tables of terms, medical concepts, medical diagnostic codes, or medical procedure

codes (HIPAA 1996). This very broad definition encompasses a wide array of different data types from demographics such as race, ethnicity, state, and country to very detailed clinical data elements such as clinical findings, procedures, or laboratory results. It is important for the array of code sets to be standardized so that data can be easily shared between organizations.

To move forward the standardization effort, the Health Information Technology Advisory Committee (HITAC) was established by the 21st Century Cures Act. This committee focuses on setting forth standards for the formation and use of federal advisory committees. HITAC recommends policies, standards, implementation specifications, and certification criteria relating to the implementation of a health information technology infrastructure to the National Coordinator for Health Information Technology (ONC n.d.). Nationally and locally, it aims to advance the electronic access, exchange, and use of health information for providers. HITAC unifies the roles of, and replaces, the Health Information Technology Policy Committee and the Health Information Technology Standards Committee that were in existence before the enactment of the 21st Century Cures Act. The 21st Century Cures Act discusses the importance of patient's access to the EHR data. In other words, if a patient request data, the Cures Act prohibits providers, health information exchange networks or facilities from preventing the access of that data. As you can imagine, this is quite complex process and there is still a lot of confusion surrounding what are best practices and how best to handle individual patient requests.

Classification Systems

As part of the standardization in healthcare, adopting a well-defined, formal system that controls the vocabulary of data would allow terms to be used consistently across healthcare. This is the purpose of classification systems, which makes it easier for healthcare entities and providers to share information with each other. Classification systems have a wide range of uses in healthcare, including the collection and reporting of public health data; the design and administration of healthcare reimbursement systems; and the development and use of quality measures and other performance measures.

International Classification of Diseases

The standard classification system for epidemiology, population health, and clinical systems in the US, and many other nations, is the International Classification of Diseases (ICD) (WHO 2022). The ICD system is used by clinicians, researchers, insurance companies, and other agencies to classify diseases and health conditions using data extracted from health records, death certificates, and other treatment and vital documents (WHO 2022). Once the diseases are classified, the data can be stored and transmitted electronically for coding, billing, reimbursement, quality management, and research. Additionally, the World Health Organization (WHO) receives ICD data from many countries that they aggregate to compile mortality and morbidity statistics (WHO 2022). Thus, ICD is used for health trend analysis globally.

Mandatory reporting of death statistics is another use of ICD data. In fact, the ICD system was originally developed by the WHO for **mortality** reporting; that is, for reporting the cause of death on death certificates (CDC 2016). The mortality data are stored in the National Vital Statistics System and are used to present the characteristics of those dying in the US, determine life expectancy, and compare mortality trends with other countries. For example, the life expectancy for the US population in 2020 was 77.0 years, a decrease of 1.8 years from 2019. The age-adjusted death rate increased by 16.8 percent from 715.2 deaths per 100,000 standard population in 2019 to 835.4 in 2020. Having ICD

data captured on the patient records was one of the reasons as to how we are able to capture the above statistics.

Additionally, with the collection of ICD coded data, healthcare professionals are able to establish and identify important public health trends that are key in many different public health administration efforts, such as creating and implementing new policies, resourse allocations, and education and training of medical staff. Some examples to this effect include the following findings from 2020:

- Nine of the top 10 leading causes of death remained the same as in 2019, with heart disease and cancer remaining as the two leading causes and COVID-19 in third place.
- The infant mortality rate decreased by 2.9 percent in 2020 with a new record low of 541.9 infant deaths per 100,000 live births.
- Almost 20,000 infants died in the United States in 2020. The five leading causes of infant death for each year is identified from capturing these types of medical codes.

ICD-10

Currently, the 10th revision of ICD is used in the US while WHO has implemented the 11th revision along with approximately 35 countries adopting it. On October 1, 2015, the United States began using the ***International Classification of Diseases, Tenth Revision, Clinical Modification*** **(ICD-10-CM)** for reporting diagnoses, inpatient procedures and services in all healthcare settings, and deaths. Similarly, the ***International Classification of Diseases, Tenth Revision, Procedure Coding System*** **(ICD-10-PCS)** provides codes for all procedures performed on inpatients. Using ICD-10-CM and ICD-10-PCS enables greater comparability between the US and other countries and, within the US, between patient mortality and morbidity codes.

ICD-10-CM includes a total of 21 chapters and a maximum of seven characters per code (see figure 3.1); the inclusion of alphabetic characters as the initial character for all codes and as a seventh character code extension when applicable; and the use of a placeholder X when necessary. Within ICD-10-CM, factors influencing health status and external causes are listed in the main classification; the sense organs and nervous system disorders are separated; injuries are grouped by body part; the excludes notes have been moved to the beginning of each chapter; and postoperative complications are in the procedure-specific body system chapters.

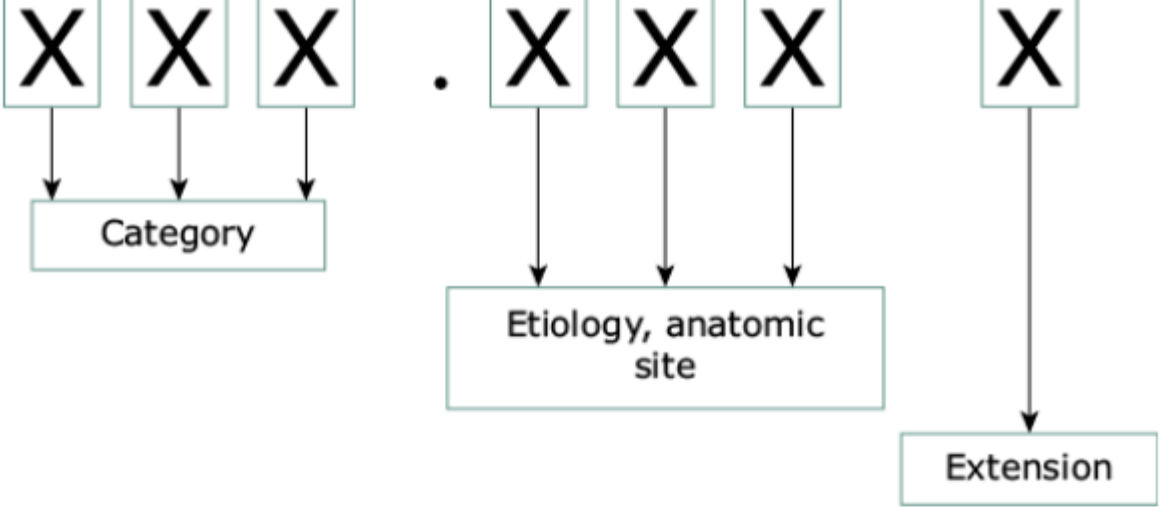


Figure 3.1. Format of ICD-10-CM

▶ Expand/collapse Extended Description

The move to ICD-10-CM was a huge transition in the US. When compared to earlier versions, ICD-10 is more specific and granular. However, this increased specificity requires more precise documentation to ensure that medical record coders can accurately assign the correct codes. After the transition, clinical documentation specialists were finding gaps in documentation that made it difficult to accurately assign ICD-10 codes (Tompkins 2016). Problem areas where coders need greater documentation specific are laterality (which side of the body), location, drug and substance abuse, and disease manifestations type, acuity, and stage (Tompkins 2016). Physician education and increased physician engagement are needed.

The Centers for Medicare and Medicaid Services (CMS) contracted with 3M Health Information Systems to develop ICD-10-PCS, a procedural coding system for inpatient procedure reporting in the United States. ICD-10-PCS has a seven-character alphanumeric structure, with each character having many different possible values, as the letters and numbers are mixed throughout the code. The first character can either be alphabetic or numeric. The digits 0 through 9 and the letters A through H, J through N, and P through Z are used. Letters O and I are omitted to avoid confusion with 0 and 1. Figure 3.2 shows the meaning of each of the seven characters from section to body system to root (or main) operation to body part to approach to device to qualifier.

To assist the transition from ICD-9 to ICD-10-CM, CMS collaborated with the CDC to generate General Equivalency Mappings (GEMs) for translating ICD-9-CM codes into ICD-10-CM/PCS codes (Bowman et al. 2016). The GEMs users include payers, programmers, researchers, and informatics professionals. GEMs are useful for retrospective data analysis and long-term studies involving data collected from both classification systems (Bowman et al. 2016).

Depending on public health needs, WHO and affiliated countries routinely monitor for the appropriateness of the implementation of the ICD codes in each country and are able to add or modify codes depending on certain situations. The recent COVID-19 pandemic was one such example where healthcare providers needed to add new codes to the classification system. At the start of the global pandemic, there was a need for documentation of the signs, symptoms, diagnosis, and procedures of the diseases. In the US, from January 1 to March 31, 2020, code B97.29 (other coronavirus as the cause of diseases classified elsewhere) was used. As of April 1, 2020, code U07.1 (2019 Novel

Coronavirus, COVID-19) is used to record COVID-19 instead. The WHO is continually monitoring and adding new codes as needed.

page 48

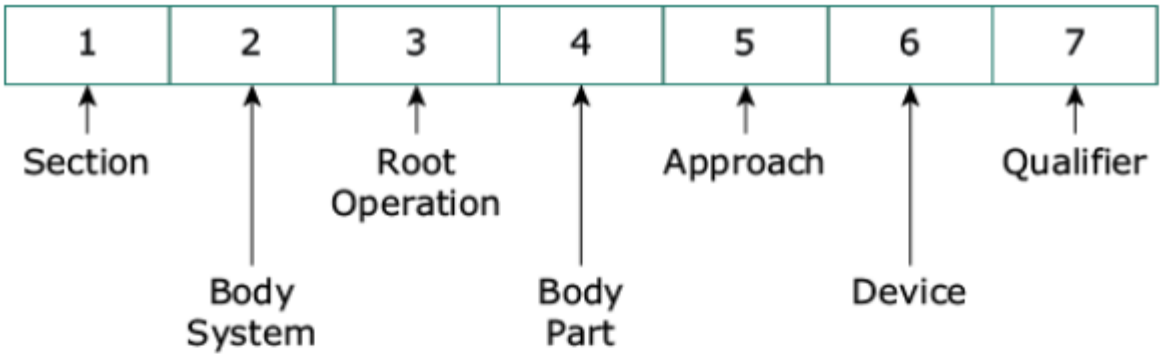


Figure 3.2. Meaning of characters for medical and surgical procedures

▶ Expand/collapse Extended Description

ICD-11

The present revision of the ICD, the 11th, was published on January 1, 2022, by the WHO in Geneva. ICD-11 promises substantial improvements over historical rendering, embracing a tripartite structure.

1. An ontological core in partnership with the Systemized Nomenclature of Medicine (SNOMED)
2. A foundation layer of richly interlinked concepts with explicit preferred terms, fully specified terms, synonyms, definitions, and detailed attributes.
3. Several linearization deriving from the Foundation Layer that will resemble the traditional mutually exclusive and exhaustive coding structure of historical ICD tabular versions. A **linearization** is "a subset of the ICD-11 foundation component, that is
   - fit for a particular purpose: reporting mortality, morbidity, or other uses
   - jointly exhaustive of the ICD universe (foundation component)
   - composed of entities that are mutually exclusive of each other

The most important linearization will be the standard derivative used for morbidity coding in hospitals and clinical care. A proper and much smaller subset of this linearization will be used for mortality coding. Distinct from these will be specific linearization for primary care, low-resource delivery environments in developing countries, and various specialty derivatives such as neurology, cancer, or psychiatry. Finally, the historical pattern of nations creating their own national modifications, such as the American ICD-10-CM (where the CM is the US Clinical Modification), can correspondingly be accommodated by nations invoking their own derivatives from the Foundation Layer of ICD-11, the advantage being that because all linearization derive from the Foundation Layer, establishing linkages and mapping between and among virtually all linearization—morbidity, mortality, specialty, primary care, or national derivations—can be achieved by mapping through the Foundation Layer in a hub and spoke model, rather than creating point-to-point mappings among the permutations.

page 49

An additional feature of ICD-11 will be its use of post-coordination to avoid the excessive repetition evident in ICD-10-CM for many injury patterns. For example, each injury has a separate code for initial encounter, subsequent encounter, and sequela. Post-coordination means that the terms with modifiers are created as they are needed rather

than having the developers of the classification try to conceive of all possible permutations for inclusion. Parameters such as severity, stage, detailed anatomy, acuity, or topography (for example, laterality, anterior, distal, and so on), can be rendered using post-coordination rather than creating hundreds of precoordinated terms for each of the combinations on such dimensions. Nevertheless, to maintain longitudinal continuity with historical versions of ICDs, the ICD-11 mortality linearization will retain the approximate level of pre-coordination evident in ICD-10.

The developers of ICD-11 recognize that post-coordination presents the possibility of multiple ways to code the same concept. An example would be post-coordinating a disease and severity when a precoordinated version of disease by severity might already exist. To avoid violating the requirement that a statistical classification must have mutually exclusive categories (to avoid counting the same things twice or separately in different categories), ICD-11 introduces a body of "sanctioning rules" that are intended be computationally executed, not manually practiced, because there will likely be many thousands of such rules. An example of such a sanctioning rule might be mapping a post-coordinated expression of renal failure and a severity code to its corresponding precoordinated version.

An interesting speculation, as one proposal to eliminate an expansion of national modifications, is to consider that some modes of pre-coordination would be required or prohibited as a set of restrictions. For example, the US might require the post-coordination of laterality where appropriate. Building on the notion that national modifications of ICD may become simply various linearization derived from the Foundation Layer of ICD-11, a country's transition to ICD-11 can be facilitated by rendering their current system, such as ICD-10-CM in the United States, as a linearization of the ICD-11 Foundation Layer (this is technically straightforward), and simply evolving that linearization to become incrementally closer to the standard ICD-11 Morbidity Linearization over time. This provides a nondisruptive pathway for graceful adoption of the new revision without the abrupt transformation witnessed around ICD-10-CM adoption in the US today.

The present revision process of the ICD, the 12th, is under way at the WHO in Geneva.

Clinical Terminologies

It is implausible that a single terminology or classification can or should function as the only reference set for all the complexities of clinical information. Laboratory or genomic data are not well suited to representation in any version of the ICD, even the ever-versatile ICD-11. Thus, many purpose-specific terminologies are widely used today. The more common terminologies used are presented below.

Healthcare Common Procedure Coding System

The **Healthcare Common Procedure Coding System (HCPCS)** consists of two levels or systems. Level I is the Current Procedural Terminology (CPT), a large, well-curated catalog of clinical and surgical procedures maintained by the American Medical Association (AMA) as a proprietary and copyrighted resource. Its primary purpose is to enable the classification of resource categories for reimbursement, although it has levels of detail and granularity that support its reuse for secondary analyses, such as quality improvement or comparative effectiveness research. Although comprehensive, it is lacking for coherent overall structure, hierarchy, or conceptual organization, rendering its facile adoption for secondary use sometimes challenging.

Level II of the HCPCS is maintained by CMS. CMS creates and administers the Level II HCPCS codes as alphanumeric identifiers for products, supplies, and services not included in CPT. The codes, which consist of a leading alphabetic character followed by four numbers, are used primarily for claims processing.

SNOMED CT

The Systemized Nomenclature of Medicine (SNOMED) developed by the College of American Pathologists merged with the Clinical Terms (CT) vocabulary from the United Kingdom to create SNOMED CT. **SNOMED CT** is a logically organized and massive terminology, mostly ordered by a formal description logic, which permits it to benefit from the classifying reasons developed by computer science and the fields of formal ontology and logic. Each entry in SNOMED CT contains a concept identifier, a description that helps define the concept and relationships or attributes for the entry. It has highly detailed and granular characterization of diseases, findings, symptoms, anatomies, manifestations, and many related axes. It is the largest, most complex, and comprehensive terminology about clinical medicine (NLM 2016). Architecturally, it is intended for representing hierarchies and relationships between concepts (Giannangelo 2015). It can be used to record smoking status and to update smoking status changes (Giannangelo 2015). Figure 3.3 is an illustration of the multiple granularities available in SNOMED CT.

SNOMED CT use in the US is growing because it is specified as an acceptable standard for several Meaningful Use (MU) requirements for entering structured data into EHRs. Examples include representing patient problem lists, symptoms, and procedures. It is also recommended, in MU criteria, for encoding lab, diagnostic, encounter, substance, family health, and medical device data (Giannangelo 2015). SNOMED CT is more easily understood if patient context data are entered into the system along with the data used for generating the SNOMED CT codes (Nandigam and Topaz 2016). However, this is time consuming for physicians. An alternative is to define a commonly used subset of SNOMED CT codes for a given setting and have providers select the subset codes that match the patient's problems (Nandigam and Topaz 2016). In fact, the National Library of Medicine (NLM) has mapping guidelines for SNOMED CT and ICD-10-CM codes that can assist informatics specialists who are designing electronic systems (NLM 2016).

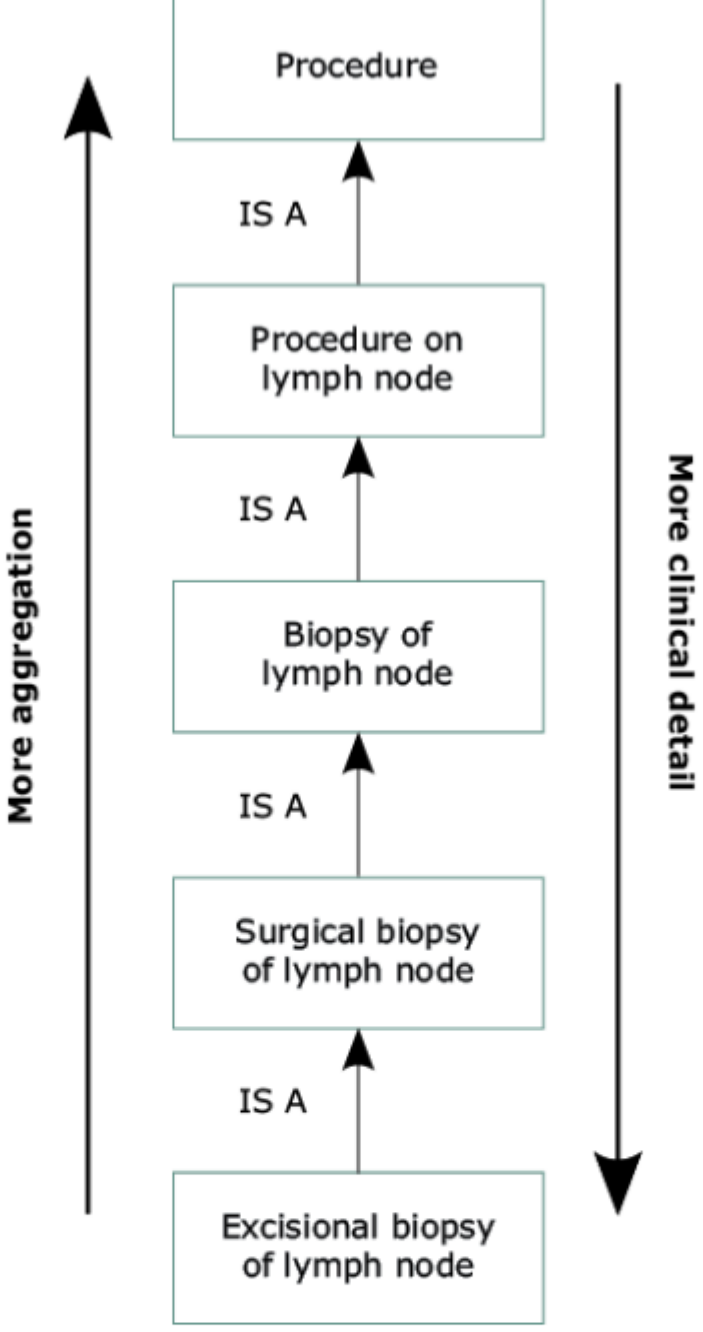


Figure 3.3. Multiple levels of granularity in SNOMED CT

▶ Expand/collapse Extended Description

Figure reprinted with permission of the International Health Terminology Standards Development Organisation (www.ihtsdo.org).

LOINC

Logical Object Identifiers Names and Codes (LOINC) is a dictionary of laboratory codes and clinical test descriptors, originated by the Regenstrief Institute and funded by federal support grants as a public resource for the common good. It is publicly usable and used worldwide without any fees or licenses. Laboratory codes are synthesized from a six-part naming model: analyze/component (for example, sodium), property (for example, concentration), timing (for example, a moment or integrated over a defined period),

sample type (for example, blood serum), scale (for example, quantitative), and optionally method. LOINC has two sections, lab and clinical codes (Oachs 2020). Clinical codes examples are vital signs, EKG, and cardiac echocardiogram (Oachs 2020). Variations on this six-part model are also used for clinical measurements (for example, blood pressure) and clinical scales or survey instruments (for example, the Glasgow Coma Scale). Selected examples from the LOINC tables are found in figure 3.4.

Figure 3.4. Selected examples from LOINC tables

LOINC #	Long Common Name	CLASS	Class Override
6019-4	Almond IgE Ab [Units/volume] in Serum	Allergy	Allergy
6020-2	Alternaria alternata IgE Ab [Units/volume] in Serum	Allergy	Allergy
19000-9	Vancomycin [Susceptibility]	ABXBACT	Antibacterial susceptibility
7059-9	Vancomycin [Susceptibility] by Gradient strip (E-test)	ABXBACT	Antibacterial susceptibility
38483-4	Creatinine [Mass/volume] in Blood	Chem	Chem
2160-0	Creatinine [Mass/volume] in Serum or Plasma	Chem	Chem
35591-7	Creatinine renal clearance predicted by Cockcroft-Gault formula	Chem	Chem

This material contains content from LOINC® (http://loinc.org). The LOINC table, LOINC codes, and LOINC panels and forms file are copyright © 1995–2022, Regenstrief Institute, Inc. and the Logical Observation Identifiers Names and Codes (LOINC) Committee and available at no cost under the license at http://loinc.org/terms-of-use.

LOINC was originally adopted for Health Level 7 (HL7) clinical messages because most labs use proprietary test codes and an exchange standard for lab results was needed. LOINC is now endorsed as the preferred coding for laboratory data, and drug label section headers (Oachs 2020) as specified by the MU regulations for the US Department of Health and Human Services (HHS; ONC 2013a). Use of LOINC facilitates data exchange and interoperability between EHRs and lab systems. However, given the structure and complexity of LOINC lab codes, converting a lab system's proprietary codes into LOINC codes is a substantial undertaking, because additional information will need to be captured for the six-part model. After the model data are entered into the local systems, the local lab result codes can be mapped to LOINC codes by using Regenstrief LOINC Mapping Assistant, a desktop program (Giannangelo 2015).

RxNORM

Another standard for electronic data exchange is **RxNorm,** a standardized nomenclature for clinical drugs created and curated by NLM (2016). RxNorm is the MU recommendation for pharmaceutical names and codes in the United States. To be certified, EHRs must satisfy the e-prescribing medications vocabulary standard, which can be accomplished with RxNorm codes (NIST 2016). Additionally, it is used in pharmacies and drug interaction software such as First Databank and MediSpan (NLM 2016). RxNorm is a relatively sophisticated database that enumerates "orderable drugs"

such as Chewable Amoxicillin Tablet, 250 mg, and links these to their component medicines (Amoxicillin), trade name variations, and active ingredients.

Data in RxNorm are standardized from names found in many other drug nomenclatures, which facilitates interoperability. A typical listing for a drug contains the ingredients, strength, actions, and dosage forms (NLM 2016). Linkages within the system can present the drug's generic and brand names, as well as the active ingredients (Oachs 2020). For example, RxNorm includes the Veteran's Health Administration National Drug File Reference Terminology used for coding drug actions and usages. The US Food and Drug Administration's free Adverse Event Reporting System (FAERS) data for drug names can be mapped to RxNorm codes (Banda et al. 2016).

HL7

The previous section emphasized terminology, vocabularies, and classifications, but illustrated only a few. Additionally, there are information model and messaging standards that provide syntax to complement the raw semantics of controlled terms, or put another way, the context for the words. A detailed discussion of these standards is beyond the scope of this chapter but includes simple structures such as HL7 Version 2 bar-delimited ASCII strings as well as the more elegant Consolidated Clinical Data Architecture specification more recently emergent from HL7 and destined for use of HIE in a meaningful manner. This section describes other emerging data and information uses.

The HL7 Clinical Information Modeling Initiative (CIMI) is an international initiative that is unifying the midlevel modeling of healthcare data structures for representing clinical observations such as blood pressure measurements (Blobel et al. 2014). One goal is to facilitate seamless data exchange of medical records and decision support tools by providing a reference model for informatics developers to use that can be translated in Unified Metalanguage (UML), a commonly used existing modeling language (Blobel et al. 2014). HL7 CIMI has been used as one of the healthcare-specific perspectives to represent data from outcome measurement (Blobel et al. 2022).

To facilitate the creation of clinical information modeling (a product of the Hl7 CIMI) and the subsequent system implementation, clinical information modeling tools (CIMTs) can be used (Moreno-Conde et al. 2015). For example, domain experts and information technologists collaborate to define clinical data models for EHRs, which are used to write the software components that are tested and implemented. Moreno-Conde et al. explored the level of consensus of the essential requirements for CIM tools by interviewing experts and synthesizing existing literature. Essential requirements included supporting semantic interoperability of EHRs, importing and exporting data in XML, supporting version management, and collaborating when creating CIMs (Moreno-Conde et al. 2015).

In 2015, the HL7 organization developed the latest version of HL7, which sometime is also considered to be the fourth version of HL7, known as the Fast Healthcare Interoperability Resources (FHIR). FHIR use a set of resources that can be combined into systems for mobile phones, cloud-based systems, and EHRs (HL7 2015). Based on JSON, a data exchange format commonly used in mobile and web development, HL7 FHIR provides a framework to allow agile development of HIE. The standard implements RESTful web services as the backbone technology to facilitate data transfer between healthcare entities (HL7 2022).

As of 2022, FHIR has not replaced HL7 or X12 standards. However, CMS is currently exploring the use of HL7 FHIR for price transparency efforts, including the process in Advanced Explanation of Benefits. This effort might promote the use of FHIR HL7 to enable payers to send explanation/educational materials to patients.

Check Your Understanding 3.1

Answer the following questions.

1. Which of the following is utilized in the US for reporting the cause of death on death certificates?
   a. ICD-10-CM
   b. SNOMED CT
   c. LOINC
   d. RxNORM
2. Which of the following have a wide range of uses in healthcare, including the development and use of quality and performance measures?
   a. Code sets
   b. Classification systems
   c. Clinical terminologies
   d. Data set standards
3. Which is not a benefit that ICD-10-CM/PCS codes provide?
   a. Allowing more comparability
   b. Providing additional detail
   c. Assisting identification of factors influencing health
   d. Require more specific documentation to assign the codes
4. Which of the following would be considered metadata?
   a. ICD-10-CM codes
   b. Source of the data
   c. Trend analysis
   d. Data of admission
5. A dictionary of terms is:
   a. Vocabulary
   b. Terminology
   c. Classification system
   d. Code set

6. Which sentence does not describe linearization?
   a. Linearization allow fitting for a particular purpose, such as reporting mortality, or morbidity
   b. Linearization is jointly exhaustive with the ICD universe (foundation component)
   c. Linearization is composed of entities that are mutually exclusive of each other
   d. Linearization is mutually exclusive with ICD-11 foundation component

Provide short answers to the questions below.

7. What is the difference between Classification Systems and Terminologies?
8. Why is the HL7 FHIR considered to be more agile to be used in development?

Health Information Exchange

Healthcare entities commonly exchange information about their patients to ensure continuity of care. By exchanging health information, healthcare entities can minimize effort related to gathering information about patient that has been previously conducted in other healthcare entities while ensuring that the patient's information is consistent across care. **Health Information Exchange (HIE)** is defined as "exchange of health

information electronically between providers and others with the same level of interoperability, such as labs and pharmacies" (AHIMA 2017). While HIE can happen directly between entities, in many situations, having an organization facilitates the exchange can ensure that the information exchanged are fast, consistent, and comprehensive with as minimal exchange as possible. For example, imagine HIE between three healthcare entities. With an organization facilitating HIE, a healthcare entity needs only to exchange the information with the organization instead of conducting it multiple times with each other entities individually. In practice, the acronym HIE is commonly used interchangeably for both the process of information exchange and the organization facilitating the exchange as most information exchange process do need to go through the organization that facilitates the exchange. Therefore, for the remainder of this chapter, the acronym HIE will be used interchangeably as well to refer both the information exchange process itself and the health information organization that facilitates the exchange of the health information. However, when the discussion focuses on the business model and structure of the organization, the term Health Information Exchange Organization (HIO) will be used to help differentiate between the two.

Since its inclusion in the provisions of the Health Information Technology for Economic and Clinical Health (HITECH) sections of the American Recovery and Reinvestment Act (ARRA), HIE has become commonplace in the US (Vest and Kash 2016). The legislation is being implemented via the CMS EHR Incentive Program, which pays providers for their meaningful use of EHR technology. The plan for meaningful use includes the articulation of three stages so providers can adopt EHRs and increase the functionality and interoperability of EHRs over time. Stage 1 of meaningful use requires providers to have the ability to perform HIE (HHS 2010). Initially, Stage 2 of meaningful use increased the requirements for HIE in multiple ways. Eligible providers needed to participate in generating and transmitting prescriptions electronically; provide patients with the online ability to view, download, and transmit their health information; submit electronic data to immunization registries; and use secure electronic messaging to communicate with patients on relevant health information (CMS 2012).

In 2015, Stage 2 and 3 requirements were streamlined to address providers' comments to remove redundant, duplicative, and topped out measures (CMS 2015b). In the revised requirement, providers can select the measures for reporting that are a better fit for their professional needs and, starting in 2015, the reporting period becomes any continuous 90 days period within the calendar year (CMS 2015b). The streamlining includes reduction to requirements, such as a reduction from 18 objectives to 10 objectives for professionals, and reduction from 20 objectives to 9 objectives for hospital (CMS 2015a). For stage 3 in 2017 and later, application program interfaces to allow increased data access and public health reporting are required to allow interoperability (CMS 2015a). Additionally, health information professionals will need to perform CMS measures reporting.

In 2018, the meaningful use programs were overhauled and renamed as the Promoting Interoperability Programs, which measure healthcare entities based on performance and results around interoperability as well as giving patients access to their data (Federal Registrar 2018). Stricter compliance approaches, such as using thresholds and check-the-box measures, were removed from the program to allow more flexibility in implementing the interoperability functions that fits with the need of care and services as well as the capacity of the system. Proposal to advance the healthcare interoperability by removing information blocking was also submitted in March 2019 and is implemented in the 21st Century Cures Act.

The Office of the National Coordinator for Health Information Technology (ONC) provided assistance for establishing HIEs with the State Health Information Exchange Cooperative Agreement Program with the goal of enhancing "states' efforts to rapidly

build capacity for exchanging health information across the health care system both within and across states" (ONC, DHHS 2012). In 2020, ONC moved their incentives from establishing HIEs to support current HIE through sustainability program fundings, such as the Strengthening the Technical Advancement and Readiness of Public Health Agencies via Health Information Exchange (STAR HIE), which focuses on reinforcing state and local HIE infrastructures to give public health agencies access to health data with additional aim to support communities hit hard by COVID-19. However, as the grant funding dwindles, HIEs struggle to be sustainable (HealthIT 2020). This section will discuss HIE; the history, challenges, and clinical and financial benefits; provider and consumer perceptions, and sustainability.

History of Health Information Exchange

HIE was conceived prior to the HITECH legislation. Networks titled Community Health Management Information Systems (CHMISs) were first established in 1990 through grants from the Hartford Foundation (Vest and Gamm 2010). CHMISs focused on the development of a centralized repository of clinical, demographic, and eligibility data. However, CHMISs experienced significant obstacles such as cost as CHMIS networks preceded the advent of widespread, reasonably priced, standards-based high-speed Internet access. As expensive hardware, software, and network connections were required for CHMISs, the lack of data standardization was an obstacle, and both patients and providers expressed privacy concerns (Vest and Gamm 2010).

In 1994, we saw the emergence of Community Health Information Networks (CHINs). These commercial initiatives focused on financial savings, with little emphasis on the public health benefits that might accrue. CHINs encountered problems when competitors limited the data sharing functionality to prevent any perceived loss of competitive advantage, the ability of vendors to charge fees, and differences in the fees charged different participants raised questions regarding any financial benefits (Vest and Gamm 2010). The main lesson from CHINs was that the benefits could not be assigned to one small group (in this instance, providers); instead, they must be spread and shared across the healthcare industry.

The late 1990s and 2000s saw the rise of the Regional Health Information Organization (RHIO) in the healthcare industry. RHIOs are the immediate precursor to the HIO. An RHIO is "a neutral, third-party organization that facilitates information exchange between providers within a geographical area" - organizations that are founded with no commercial interest in the use of the healthcare data itself, which would focus in facilitating information exchange between providers (Vest and Gamm 2010). The obstacles to RHIO success were in identifying a sustainable business model that allows them to maintain the technology infrastructure, assuring privacy and security of the data, while overcoming the issue of competitor distrust. Since many RHIOs and HIOs were often created through (and operate) from a grant, these same problems exist for HIOs today when the funding dries up.

Now, however, the HITECH legislation and MU regulations mandate HIE functionality and its use. Researchers have proposed the following considerations for the continued development of US HIEs:

- Adopting an improved business model, perhaps with the increased use of quality incentives
- Acknowledging the public health benefits of health information, which might include improved registries for chronic conditions and much better method to develop

effective prevention and operational capabilities for detecting and countering biological threats.

- Monitoring HIE activities so both providers and consumers trust that health information is private and secure
- Setting state borders as the standard geographic unit for HIEs, which now is operationalized by ONC (Vest and Gamm 2010)
- Implementing strategies and features that address barriers to access and value-added services to forward adoptions of user-initiated HIEs (Rahurkar et al. 2021)

From 2015 to 2017, the CMS Health Information Exchange Objective for EHR incentive required that a summary of care record be sent by all eligible providers when a patient is transferred to another provider or care setting (CMS 2015c). The care record must use certified EHR technology (CEHRT) and must be transmitted for more than 10 percent of patient care transfers or referrals (CMS 2015c). Summary of care documents must include the current problem, medication, and allergy list. Additional fields are name, provider, procedure, lab, demographic, and care plan items (CMS 2015b). These requirements are still valid even after the transformation of the MU program into Promoting Interoperability Programs in 2018.

Structure and Adoption

HIEs can have different technical and organizational structures, which is why they are separated from HIOs. Whereas HIEs refer more to the process of interoperability, and usually focus on the technical aspect of the exchange, HIOs refer to the organization facilitating the exchange and usually focus on the organizational structures and policies surrounding the exchange. Given the history, not one superior structure has emerged. It is reasonable to presume that HIO operations will become standardized in the future; however, the current industry is still evolving in regard to finding the best HIOs structure and policies.

Technical Structure

There are two main technological approaches to HIE. The first is an **aggregated HIE**. The aggregated HIE combines all of the data into a centralized repository with a master patient index (MPI) or record locator service. The different HIE participants query, or send a request to, the repository to obtain demographic, clinical, or other information (Hess 2011). The aggregated HIE might be preferred by persons or local governments trying to maximize the public health benefits of HIEs; however, they present greater risks to the privacy and security of data. All of the data, some of it very detailed and potentially sensitive, must flow to the centralized repository for aggregation process to be conducted.

The second type of HIE model is a **federated HIE**. Federated HIEs are designed as provider-to-provider networks using the internet for connectivity, with no central repository of data (Hess 2011). A central entity maintains the MPI or record locator service, which is used to determine where patients have been treated previously and where their health information might exist. Providers send requests for data information directly to other providers. This type of arrangement might require an additional step to retrieve the data (as the data is not stored in a central location); however, given the distributed maintenance of the data, it is viewed as having stronger privacy and security protection.

Organizational Structure

The technical structure of the HIE is separate from the organizational structure of the HIO. HIOs tend to be one of three types: public, cooperative, or private.

Public HIOs are entirely supported by state government agencies or are semi-independent, with governmental financial backing (Hess 2011). There is some consideration of HIOs as a public good, such as the road system (Vest and Gamm 2010). The case for the long-term governmental support of HIOs is, unfortunately, not an easy one to make. The occurrence of natural disasters supports the public health benefit claims. Examples include Hurricane Sandy in 2012 at New Jersey where many were relocated for an extended period of time. Hurricane Sandy also significantly impacted New York City with hurricane damage to many structures including healthcare facilities, evacuations of patients and others, and loss of power and public transportation for an extended period of time affecting millions of people. With many displaced population, public HIOs would allow health and healthcare data to follow the individuals and still accessible and can be used even when they are not currently located in their original area of service.

Cooperative HIOs arise when relationships are formed between what are otherwise competing healthcare providers for the purpose of exchanging information (Hess 2011). Sometimes these HIOs struggle with financial viability and sometimes they are quite successful. Often the difference rests on the community or geographic region's concerns. For example, hospitals in Austin, Texas, came together to form the Integrated Care Collaborative (ICC) in 1997. The original goal was to address mutual concerns related to low-income and uninsured residents excessive use of emergency rooms (ERs) (ICare 2022). Because all the hospitals with ERs within the region could easily see potential financial advantages, hospital administrations agreed to contribute funds to establish and maintain the HIO.

The third HIO structure is the **private HIO**. This type involves a single, integrated delivery system (IDS) that connects all of the different providers with a healthcare entity. One example is the Veterans Health Administration (VHA). For years, VHA had aggregated their patient data; however, as the World War II veterans aged and began retiring or spending large parts of each year in different regions of the country, the need to transmit or exchange the data became a high priority (Tsai and Rosenheck 2012, Donahue 2018). Over a decade or more VHA established the data and other standards that allowed them to be a pioneer in health information exchange. They also undertook efforts to include non-VHA data from non-VHA providers in their health information exchange efforts (Tsai and Rosenheck 2012, Donahue 2018). This resulted in the ability to provide longitudinal picture of veterans' health record whenever and wherever a veteran receives care. In 2018, VHA, DOD, and the US Coast Guard began the process of moving all health-related data to the Genesis system, a new integrated EHR system based on CERNER's platform, further strengthening the integration and interoperability of health data between the three agencies.

Patient's Consent Option

Additionally, one very important aspect of structure and management for an HIO to be adopted by patients and consumers is the consent option chosen as part of the HIO. Each HIO must determine how patients and consumers will consent or not consent to have their data and information transmitted by, or included in, HIE operations. A white paper prepared for the Texas Health Services Authority describes five generally accepted models for defining patient consent to participate in an HIO (Gray 2011). The main points for each option are described in figure 3.5 below.

Figure 3.5. Health information exchange consent options

OPTION	PROS	CONS
No-consent model: does not require any agreement on the part of the patient Unknown: Notice requirements to patients; interface with an HIE requiring consent	Easier to administer; addresses provider concerns about incomplete information	Lack of patient control
Opt-out: a predetermined set of data are automatically included, unless the patient "opts-out" or denies access to the information.	Automatic inclusion of data for more complete data for providers, yet patients can restrict use	Procedures for documenting opt-out so providers know about it are unclear
Opt-out with exceptions: patient can choose level of participation. • Selective exclusion of personal health information (PHI) • Limited to certain providers • Limited to exchange for certain purposes	Much more patient control	Very, very difficult to implement
Opt-in: patients choose to include their data in an HIE. It is an all-in or all-out choice.	Patient has complete, up-front control; provides an opportunity to educate patients; may provide better record matching	Obtaining and communicating consent; questions as to how consent be revoked or withdrawn
Opt-in with restrictions: patient agrees up front, but can determine which data is shared, which providers can access it, or the purposes for which it can be accessed.	Patient control up front and very detailed	Very, very difficult to manage

The first consent model, the **no-consent model** does not require any agreement on the part of the patient to participate in an HIE. HIOs that have adopted this model operate in states that explicitly allow this model via legislation, such as Delaware Health Information Network Regulations on Use of Clinical Data for Approved Analytic Purposes (Gray 2011). The major benefit of this option is that providers can be more certain that any data and information they receive is complete. This ensures they are not trying to make clinical decisions using incomplete data. However, this consent option does not mean patients have no rights. The HIOs using this option usually allow patients to deny access to their data and information even if it is technically included in the HIE. This is the easiest consent option to administer as patients and consumers must take action to deny access. Providers must include the fact that they participate in an HIE in their Health Insurance Portability and Accountability Act (HIPAA) Notice of Privacy Practices that patients receive at their first visit (Gray 2011). Failure to include the statement may expose healthcare entities and providers to liability under the HIPAA Privacy Rule.

Another option is the **opt-out model**, which allows for a predetermined set of data to be automatically included in an HIE, but a patient may still deny access to information in the exchange. This model allows consumers to prevent their data and information from inclusion in the HIE. Inclusion or exclusion is all or nothing for this option; that is, either

all of the patient's information is included or it is all excluded. Issues to address in this model include the following:

- Who will collect the necessary information for patients to opt out?
- How will the opt-out information be communicated to the HIO and to other providers?
- What processes can be implemented to handle cases where a patient changes their mind? (Gray 2011)

The third **opt-out with exceptions model** makes the patient's information available in the exchange, but "enables the patient to selectively exclude data from an HIE, limit information to specific providers, or limit exchange of information to exchange for specific purposes" (Gray 2011). Although this option is understandably attractive to patients, it is not yet clear how to implement this model from a technical standpoint; that is, it is very resource intensive. As well as raising questions about the completeness of the data, since this option is basically the same with the regular opt-out model, this option inherits all challenges from the opt-out model (Gray 2011).

The fourth **opt-in model** requires patients to specifically affirm their desire to have their data made available for exchange within an HIO. This option provides up-front control for patients because their data cannot be included unless they have agreed. Disadvantages of the opt-in model are primarily the administrative complexity of implementing such an approach. Issues the HIO must address include the following:

- Who will obtain the consent?
- Will one consent suffice for participation in the system as a whole or must each provider obtain consent for his or her own patients?
- What process will be used to communicate the patient's consent?
- What happens if a patient wants to withdraw consent? Can that patient's data be removed from the HIO or will the withdrawal of consent only be effective for data developed after the withdrawal of consent was given? (Gray 2011)

Finally, the **opt-in with restrictions model** allows patients to restrict how much data and what data are available. They may enable access for certain doctors or only for certain data elements or specific usage (Gray 2011). For example, a pregnant patient can configure so only ob-gyns can access their maternity data but only for the purpose of providing warning if they have a concern on the pregnancy progress.

As of this writing there is no single mandated consent option for HIE implementation. It is imperative that HIOs choose wisely in accordance with the need of information exchange and the level of awareness of HIOs' stakeholders toward the importance of HIE as the ease of provider implementation, as well as consumer confidence in the HIO, may be impacted by the option chosen.

Adoption

The adoption of HIEs is a part of the national initiatives to help healthcare adopt twenty-first-century information technology. Specifically, the HITECH legislation provided funding for the State Health Information Exchange Cooperative Agreement Program (ONC, DHHS 2013). This section explores levels of adoption of HIEs, how HIEs are used, as well as costs, benefits, and evaluation, concluding with the expected interaction between the states and the federal HIE.

Research reveals increasing trend of HIE adoption in US. In 2021, HIE adoption has grown towards 50 percent across all US hospitals (Apathy et al. 2021). This number has

grown from around 30 percent adoption during the earlier introduction of the meaningful use program in 2014 (McCann 2014). Medicaid's shift to managed and accountable care payment systems is predicted to create a need for the type of coordinate care that is facilitated by HIOs. Lack of standards, lack of interoperability, and poorly designed and documented methods to connect with health information technologies in other healthcare entities continue to be barriers (McCann 2014). Rural hospitals are lagging in adoption of HIE due to cost constraints and staffing shortfalls (Kim et al. 2016). These and other surveys of operational HIEs reveal a continued concern regarding sustainability and financial viability, with questions about clinical benefits remaining.

Overall, HIEs at the federal and state levels focus on a public–private collaborative structure. Although the National Health Information Network was established and supported by the ONC, management of national HIE efforts was transferred to Healtheway, Inc. on October 10, 2012. Healtheway, Inc. states it "is a community of public and private exchange partners who share information based upon a shared set of 'rules of the road' and a trust agreement, called the Data Use and Reciprocal Support Agreement (DURSA)" (Healtheway, Inc. 2012). State HIEs were invited to become members of Healtheway, Inc., which established and supported a portfolio of nationwide health information network standards, services, and policies (Healtheway, Inc. 2012). In 2015, Healtheway was rebranded to become eHealth Exchange, which continues as the largest healthcare information network in the US, currently active in all 50 states.

The fact that HIEs and interoperability have been mandated as a significant part of meaningful use and are supported by funding does not provide one with any information regarding usage of the HIE and resulting financial or clinical benefits. Usage refers to the manner in which providers utilize HIE. It is important to understand how HIEs are used within clinician workflow to continue to improve functionality and value. The most common type of HIE use is based on encounters between providers and patients (Vest et al. 2012). Studies have shown HIEs are often queried for information when patients are seen in clinics and report a history of other clinical care such as a physician visit or hospital stay (Frisse et al. 2012; Vest et al. 2012; Apathy et al. 2021). These studies reported unanticipated impacts on workflow. For example, the Frisse research team found clinicians reporting additional work when sorting through their clinical data, the data from the HIE, and any paper forms (Frisse et al. 2012). Organizational factors rather than the HIE technology may affect use. For example, if the emergency department is not fully staffed, the personnel might not take the time to query the HIE. Study also found that usage varied by user and workplace setting. Physicians and others under significant time pressures used the system less than nurses, public health, and other workers (Vest and Jasperson 2012). As with other new technologies, effective incorporation of HIE data and the processes for health information exchange into the delivery of high-quality healthcare will need to develop over time.

Evaluation of the use and utility of HIEs is imperative as the country continues to implement them. One researcher who performed a meta-analysis summarized his findings with the statement, "The extent of clinical HIE benefits to date is dependent upon the context of the HIE" (Joshi 2011). For example, a community such as Baltimore, MD might reap significant benefits by tracking patient usage of ERs via HIE, leading the physicians in the ER to actually see an increase in their work or, at the least, experience a potential delay in the timing of care delivery. Likewise, the benefit of the surveillance of epidemic or outbreaks is related to the size of the exchange system (Joshi 2011). The use of HIEs become prominent during the COVID-19 pandemic in providing information on vaccination and COVID testing availability, exchanging information on public health immunization, and supporting agencies and healthcare entities via a public health surveillance dashboard that can be generated in real time to detect local community

outbreaks (Dixon 2021). HIEs can be used to improve patient care or, adversely, they can have little impact on it. Implementation and context make a difference.

These evaluations on HIE use might help justify the implementation of HIE because justifying HIE through analyzing the relationship between cost savings and utilization of HIE is difficult. Many cost savings from HIE systems do not always come back to the payers. One study did quantify an annual savings of $1.07 million in an emergency department utilizing HIE (Frisse et al. 2012). This is particularly noteworthy because a related study also points out that the perverse incentives in the US healthcare industry payment models have providers paying for HIT implementation and health information exchange efforts, while patients and payers reap the efficiency (financial) and quality benefits (Pevnick et al. 2012). This situation shows that while the society benefits from HIE, healthcare entities themselves do not see the financial benefit from HIE. Development and evolution of HIEs is likely to continue for the foreseeable future. The final structures, technical and organizational, along with ultimate clinical and financial benefits, remain to be determined by the stakeholders of HIE.

Stakeholder Perceptions

The ultimate success or failure of HIOs is dependent on two important factors: (1) providers or users accessing the HIO and using the data and information; and (2) patients or consumers allowing their data and information to be transmitted across the HIO. This section explores these factors.

Providers or Users

HIOs do not operate in a vacuum and must be supported by the healthcare organizations in a geographic region for success. Aside from the national mandate, HIEs will be more successful if the organizations believe they can bring value to themselves and their constituents. The value might be in the form of financial benefits, such as reduced duplicative laboratory testing, or by enhanced organizational reputation for technology use that contributes to the quality of patient care. Conversely, the organization may have concerns such as legal liability (for example, situation where the information shared leads to a malpractice) or loss of competitive advantage as health data are considered to be valuable (sharing of information diminish the perceived value of the information).

Within each organization participating in an HIO, users must access HIE for the benefits to be achieved. As described previously, HIE can be implemented as a central data repository of existing patient data or an index linked to historical patient data. Thus, any provider wishing to access data would be burdened with a process to request the data from HIE. It is important to understand how users, physicians, and nurses who use HIE in their day-to-day work, feel about the burden of requesting data from HIE to ensure adoption on top of their already busy activities.

Clinicians have reported perception of significant positive value from HIEs. In particular, clinicians noted benefits in using information gained through accessing HIE, including better decision making, improved treatment planning, enhanced medication tracking, and an enriched physician/patient relationship (Hincapie et al. 2011). Given that the process of HIE and HIOs are still evolving, many do not transmit or share complete EHR data, due primarily to interoperability or inability to use the data by the receiver of the information. A consistent clinician request from multiple sources was for a more comprehensive information about patient (Gadd et al. 2011; Hincapie et al. 2011, Chandrasekaran 2021). One study focused on analyzing usage to evaluate HIE utility. As might be expected with most systems, usage ranged from minimal to intensive, sometimes depending upon the setting and other times depending upon the type of user

(Vest and Jasperson 2012). The main take-away is that HIEs are similar to other health information technology, such as EHR and CPOE and will require human factors analysis and studies of usability to maximize utility.

The perceptions of benefit by organizational and provider participants in HIOs are important and will play a vital role in their adoption of HIE. The potential benefits in terms of patient health management are expected to accrue to society or geographic regions. More research to show evidence of benefits from information exchange are needed to show the true extent of HIE use as most studies focus on implementation of HIE and have not yet investigated the quality of HIE and its impacts.

Consumer Perceptions

Consumers and patients must agree to participate and have their data included in HIEs in order for them to be successful. In some ways it seems odd that this might even be a concern. After all, many consumers use social media, such as Facebook, to share intimate personal information. Similarly, the use of mobile banking and online shopping, as well as use other advanced technologies, put personal data out on the Internet. HHS ONC guidance states that the consumer may request the method of provider communicates using a variety of methods such unencrypted e-mail or texting. However, health data and information are sensitive, especially if they are related to conditions, such as substance abuse, addiction, or sexually transmitted diseases. That said, easing access to health information or moving it across cities, regions, and countries could be of significant benefit to patients.

To help determining what data should or should not be moved with the patient, a **data segmentation** process can be used to tag patient information so that selected portions of the record can be shared (CMS 2015d). Sensitive data like substance abuse treatments can be held private while other data such as allergy is shared. Patients can receive alerts if unauthorized providers try to access data tagged as having re-disclosure restrictions (CMS 2015d). Patients may be more likely to seek treatment for AIDs, mental health, or other socially sensitive health problems if they are confident that their data are protected to keep their privacy. HITECH notes that the Health Information Technology Policy Committee (HITPC) recommends data segmentation and suggests furthers studies to assess this methodology (CMS 2015d). HL7 has international standards for segmentation of sensitive health information (CMS 2015e).

As stated previously, the Stage 2 meaningful use standards have an increased emphasis on engaging consumers in their healthcare, with criteria calling for providers to offer patients the ability to view, download, and transmit their health information online by using secure electronic messaging to communicate with patients on relevant health information (CMS 2012). The most common implementation is to deploy a patient-facing portal that allows patient to login to a secure website to access their information. These requirements are consistent with findings that 51 percent of consumers trust their physicians' practice over a health plan, hospital, or the government to regulate the privacy and security of their electronic health information (Dhopeshwarkar et al. 2012). In a different study more than 75 percent indicated they might use PHR services to manage their electronic information with more than 80 percent supporting the electronic sharing of health information via an HIE (Patel et al. 2012). Between 78 and 80 percent of cancer patients, who often require sustained continuing care, perceived patient information access to their own records, as well as the sharing of information between providers, to be important functionality for EHRs (Beckjord et al. 2011). The consistent, most important concern for these various consumers was the privacy and security of their data and information (Dhopeshwarkar et al. 2012; Beckjord et al. 2011; Dimitropoulos et al. 2011;

Patel et al. 2012). These findings demonstrate that consumers understand the benefits and are willing to be involved if the tools are available.

Future Efforts

The ONC has focused on HIO efforts to increase system interoperability and add public health measures. Many HIOs are not able to provide complete, accurate point-of-care data due to practice such as information blocking or limited technical capacity of some EHR to exchange health data. Consumer-mediated exchange, where consumers are empowered to manage their own personal health information, might seem easy to achieve in the digital age, but in reality, hybrid records and lack of interoperability pose significant challenges to consumers wishing to assume more control over their health data exchange.

There are several barriers to consumer engagement in HIOs that are being addressed for future efforts. Historically, patients had been conditioned to see the physician as an authority figure and might be intimidated about asking to control their own healthcare data (ONC 2012). Many patients are not aware that physicians do not readily share information among themselves for patient care. Conversely, physicians are concerned that giving patients access to too much information could cause unnecessary concerns, increase office phone calls, result in misunderstandings about chart entries, or even invite litigation. Moreover, consumers who attempt to collect their data from the HIO and place it in their PHR can face numerous technical challenges, such as having no customer-facing tools to extract data directly from the exchange system and difficulty to authenticate direct access to their own health data. Limitation in authentication is also the reason why tethered PHRs connected to hospital portals only provide data from that particular hospital's portal. Entering data into untethered PHRs may be a manual process, as these systems typically do not accurately parse and import clinical data. Neither of these PHRs will provide a seamless data exchange with other healthcare systems, such as a physician's office EHR. It follows that consumers who are not tech savvy might feel that the effort of managing their personal health information outweighs the perceived benefits.

Another challenge for the future efforts lies in improving the consistency of data coding that are being exchanged in HIE. ONC's report indicated that 70 percent of all medical decisions are affected by clinical lab results and thus consistent usage of coding system can ensure appropriate information is used in decision making (ONC 2013b). The report highlights three successful approaches to consistency of coding that are used by these providers. The first approach is enforcing the use of the same coding system for data exchange between the lab's laboratory information system (LIS) and the provider's electronic system. With this approach, lab results are electronically transmitted to the providers' system. For example, this approach is used for 98 percent of the Delaware Health Information Network's providers (ONC 2013b). The second approach is storing lab results in a repository for later reporting. For example, during a typical 10-day period, the Maryland Chesapeake Regional Information System for Our Patients receives lab results from 32 to 46 acute care hospitals, which are then stored and harmonized for later reporting (ONC 2016). The third approach comes from companies providing services to HIEs for translating local lab codes to the universal standard LOINC codes and providing outreach and training to providers. Future efforts may also see other service providers conduct workshops and arrange outreach activities to introduce a standardized coding system. This will, in theory, changes the practice to be closer to the first success story but allowing new healthcare entity to easily join the exchange because of the standardized coding system (instead of potentially proprietary coding system which is being used in the first approach).

While the HITECH Act of 2009 provided funding for HIEs in 2015, the ultimate goal was to keep the HIOs sustainable financially while operating effectively (Khurshid et al. 2015). To justify adoption of HIE, healthcare entities usually need to be able to show that they can recoup the funding and resources that they invested into the effort (in other words, the effort generates either no loss or positive return of investment). Results on the return on investment (ROI) of HIEs are mixed. Khurshid's study reported that 66 percent of the respondents indicated their HIEs had positive ROIs, and 33 percent did not report positive ROIs (Khurshid et al. 2015). Unfortunately, only one-fourth of the respondents indicated that they used metrics to calculate their HIEs' investment return. The ROI metrics used included savings in labor, reducing duplication of tests and procedures, and improving communication, coordination, and health outcomes (Khurshid et al. 2015). These findings shows that ROI calculation is not common across healthcare entities that utilizes HIE. A related article from Adler-Milstein exploring the decline and the future of HIE and HIOs' efforts reported a total of 106 HIOs across the nation in 2016, down 11 percent from 119 in 2012 (Adler-Milstein et al. 2016). Among these HIOs, approximately 50 percent indicated they were sustainable, an increase from 24 percent reported in 2012, and an additional 9 percent covered most of their costs (Adler-Milstein et al. 2016). Unfortunately, this trend persists as these investigators' updated findings from a study in 2021 identified a reduction of around 20 percent of operational HIOs (Adler-Milsten 2021). These financial concerns about sustainability raise the question of whether states and individual communities can take on the burden of supporting the HIO's multimillion dollar operational costs. HIOs seeking funding would be advised to expand their business model by offering paid services beyond data transmission, for example by providing big data analytics and artificial intelligence/machine learning based decision support services (Adler-Milstein et al. 2016; Adler-Milstein et al. 2021).

Check Your Understanding 3.2

Match the HIE consent terms with the appropriate descriptions.

a. No consent
b. Opt-in
c. Opt-in with exceptions
d. Opt-out
e. Opt-out with exceptions

1. ___Data are automatically included, though the patient can limit the data to specific types.
2. ___Does not require any agreement on the part of the patient.
3. ___Data are automatically included unless the patient chooses not to participate at all.
4. ___Patient agrees up front but can limit the data included.
5. ___Patients choose to include their data in an HIO. It is an all-in or all-out choice.

Choose the correct answer.

6. Which of the following statement is true about meaningful use?
   a. Stage 1 meaningful use focus on codification of patient information from a digital format into a smart format
   b. Stage 2 meaningful use focus on recording references from primary practices to specialties
   c. Stage 3 meaningful use characteristics include more focus on patient engagement.

d. Stage 4 meaningful use focus on advanced algorithm using artificial intelligence to improve public health
e. Stage 5 meaningful use focus on ensuring privacy of personal health information

7. Which is not an issue commonly associated with HIE adoption?
   a. Lack of consensus on standards
   b. Poorly designed interfaces
   c. Limited return-of-investment evidence
   d. Information blocking practice
   e. Built-in interoperability features in EHR
8. What is considered one benefit of HIE?
   a. Providing new business model that provides better context of each patient for advertisement purpose
   b. Improving public health through development of registries for chronic conditions and much better method to develop effective prevention and operational capabilities for detecting and countering biological threats.
   c. Protecting healthcare entities from liability
   d. Limiting practice to only within a state border
   e. Implementing strategies to increase return-of-investment via out-of-state service fee
9. Who are not considered as stakeholder that can access HIE?
   a. Doctors
   b. Pharmacists
   c. Nurses
   d. Patients
   e. Informal Caregivers

Interoperability

Interoperability is commonly known as the ability of different systems to connect and communicate in a coordinated way. In health and healthcare, this ability allows data of patient to flow between healthcare entities and follow the patient as they move throughout the continuity of care across multiple healthcare entities. Interoperability allows lab results, diagnoses, allergies, prescribed medications, and many vital health information to be accessible by care providers without conducting more tests and assessments, thus making healthcare more efficient and consistent. This concept is the backbone behind HIE which enables the movement of health information between providers that may benefit patients as they seek and receive care in a variety of settings. Interoperability also allows health data to be moved to central locations, or in other words, to be aggregated, for purposes beyond care, such as for public health monitoring or development of evidence to support best practice in care. As we have discussed the use of interoperability in our discussion of HIE in previous sections of this chapter, in this section, we will discusses more about the use of interoperability concept for aggregation of patient health data as well as information for public and population health.

Interoperability for Population Health

Before delving into interoperability for population health, let's look at the definition of population health itself. **Population health** can be defined as the level and distribution of disease, functional status, and well-being of a defined group of people with specified

characteristics (Friedman and Parrish 2010). The specified characteristics might be treatment by a given physician organization or admission to a given hospital with a defined time frame (Friedman and Parrish 2010). This differs from public health because public health usually refers to all inhabitants of a given neighborhood, city, region, state, or country. Population health is not usually charged with responsibility for vital statistics but focuses instead on the overall well-being of a specified group of people.

page 70

Due to the aging population in US, population health initiatives are being pursued by healthcare organizations to meet the goals of the Medicare Accountable Care Organization regulations (HHS CMS 2011). **Accountable Care Organization (ACO)** contracts issued under these regulations attempt to provide financial incentives for reductions in unnecessary treatment, such as procedures and hospitalizations, while ensuring improvement of outcomes related to quality of care (Gourevitch et al. 2012). In a very simplified definition, ACO is groups healthcare providers who come together voluntarily to give coordinated high-quality care to their Medicare patients. More in-depth discussion on ACO is available in Chapter 16. Research shows that many diseases result from a combination of individual behavioral factors, such as smoking and physical activity; the environment, including pollution and access to healthy food; and social factors such as housing and education (IOM 2001, Hardcastle et al. 2011, Bedson 2021). Prevention strategies, such as smoking cessation and the promotion of physical activity, within a given population can measurably improve its overall health. The ACO regulations reward providers financially for helping their populations make healthy choices while reducing the use of healthcare services (Gourevitch et al. 2012).

Healthcare providers' value-based care initiatives, such as reducing the number of emergency department visits and the number of readmissions, are goals of population health management systems (Yeager 2016). Data from many sources, such as the physicians' quality reporting system, clinical EHRs, or insurance company databases, are aggregated and analyzed to determine the best way to individualize preventive interventions for reducing the number of hospital admissions and emergency department visits (Yeager 2016). Benchmarks resulting from analyzing registries for diabetes or Congestive Heart Failure (CHF) can provide guidance for population health goal setting (Yeager 2016). Interfaces must be in place to facilitate data flow between disparate systems. Patients should be surveyed to determine their healthy behaviors at home (for example, diet, exercise, medication compliance). ROI calculations and monitoring care outcomes against benchmarks are essential to demonstrating population health preventive program success (Yeager 2016).

In 2017, the Merit-Based Incentive System (MIPS) for Medicare Part B was introduced to be used by providers (HcPro 2016). MIPS consolidated three programs including the Physician Quality Program Reporting System (PQRS) and the Value Based Modifier for PQRS and Medicare into a single 0 to 100-point value-based performance scale with four categories (HcPro 2016). MIPS replaced meaningful use for physicians with a merit-based system focused on patient care using interrelated technologies. Medicare reimbursement adjustments began under the MIPS system in 2019 (HcPro 2016). This transformation was supported by the Promoting Interoperability Programs, announced in 2018 and started implementation in 2020 (CMS 2022).

Ambulatory care settings present a challenge in staffing and managing the interdisciplinary care teams necessary to meet an ACO's needs for providing collaborative care and reporting on care outcomes (Haas 2016). In this setting, staffing complexities arise because the documentation is encounter based, patients are seen for care a few hours a week in a clinic, and healthcare providers direct the patient and the patient's family in the provisioning of care (Haas 2016). When providing ambulatory care, the caregivers might need to become involved in helping patients secure

page 71

transportation to appointments or get in for a visit sooner to prevent a hospital visit. Healthcare providers in ambulatory clinics might need to provide medication education with scheduled follow-up phone calls to encourage compliance with medication schedules. Poor health literacy, patient isolation, and physical limitations are also challenges to providing effective care for the ambulatory population due to the patient's perceiving their health conditions risk as low and would not impact their livelihood (Haas 2016).

To solve the challenges of ambulatory setting, community-based programs located close to patients can assist in providing education and initial counseling efforts. For example, community-based preventive interventions for cardiovascular disease and cancer prevention focused on health education, counseling, and web navigation help were determined to be a cost-effective method for promoting preventive healthcare in vulnerable populations like the poor, elderly, children, and minorities (Kyounghae et al. 2016). Success was measured as lower blood pressure, glucose, and weight, increased screening for cancers, and increased levels of exercise. Examples of specific interventions were medication safety for the elderly, bilingual workshops on the importance of cancer screening, and diabetes self-management weekly meetings (Kyounghae et al. 2016; LaMonica et al. 2021).

Investigating these scenarios are important for health information management (HIM) and health informatics (HI) profession because HIM and HI practitioners are likely to be the one charged with ensuring that interoperability allows ACOs and value-based systems to aggregate and use the healthcare data as needed. This might mean activities such as establishing a centralized data warehouse or a virtual data locator service; ensuring that all necessary data elements are standardized; protecting the privacy and security of the aggregated patient data; assuring the quality of the data; and analyzing as well as presenting the patient data to clinicians who provide care and make organizational decisions. In short, aggregated patient information is vital to population health and the success of the CMS ACOs.

Interoperability for Public Health

Public health exists at many different levels in healthcare. It is defined as the study of the health of populations in large areas like states or countries (AHIMA 2017). As has been discussed earlier in this chapter, internationally, WHO is responsible to coordinate the development and introduce the ICD globally to track the causes of mortality and morbidity across the world. This is not as straightforward as it might seem because each year sees new strains of influenza emerge, for example, as well as other healthcare crises, such as HIV and, more recently, COVID-19. Additionally, fast-paced travel, both local and national, helps spread germs and infections across the globe in a very short time. For example, Ebola hemorrhagic fever is a deadly disease, found primarily in Africa, that has been spread to the US by people who were infected. Ebola is transmitted to humans from wild animals. It then spreads among people through direct contact with the body fluids (for example, blood or sweat) of the infected person (CDC 2015). In 2014, the first case of Ebola in the US occurred in New York City when a medical aid worker who had traveled to Guinea was diagnosed (CDC 2015). Healthcare providers are at risk of getting Ebola especially when protective gloves and gowns are not worn, there is only a single approved vaccine for a specific type of Ebola virus, and the recovery period can last several years (CDC 2015). Another recent event is the spread of COVID-19, which was first detected in Wuhan, China, in late 2019, and was confirmed in the US in January 2020 (CDC 2022a). From these scenarios, we saw the need to ensure that information

about fast-paced diseases can travel even faster to ensure every nation has the awareness and can set up the capacity to mitigate and protect their population.

Each country performs its own public health surveillance, which is usually required by law. Individual countries, states, and other larger communities also have public health systems, which might have separate requirements. Generally, the requirements include a predefined list of notifiable diseases or treatments, such as immunizations, that healthcare providers must report to healthcare authorities. A list of notifiable diseases, conditions, treatments, or laboratory results required to be reported to the CDC are accessible on the CDC website, specifically in the National Notifiable Diseases Surveillance System (CDC 2022b). Examples of the information in this database include diseases that require immediate, extremely urgent notification, such as anthrax and smallpox. Then there are those that require immediate (Ebola), urgent (diphtheria and measles), and standard notification (botulism and hepatitis). For each condition, further guidance is provided by CDC for the cases in these disease categories that require notification by indicating the time frame of the condition with statements, such as confirmed and probable cause, confirmed cause, all cases prior to classification, and so on. This reporting is essential to maintaining the public health whether through outbreaks of diseases such as COVID-19, H1N1 (avian) flu, or fungal meningitis contracted by patients across the country after being treated with tainted medicine. Clear instructions of reporting mechanism would allow information to be aggregated immediately and propagated as soon as indication of fast-paced diseases are detected.

Historically, public health data reporting was done via entirely separate systems, such as hospitals reporting data with one format using one system, while physicians reported data using entirely separate methods (Shapiro et al. 2011). Eleven potential use cases have been identified where HIE can be used to improve public health (Shapiro et al. 2011). Each type of use case is listed below, along with more explanation of the benefits of HIE and HIOs.

1. Mandated reporting of laboratory diagnoses: The cornerstone of public health surveillance, this reporting is resource-intensive when performed manually. Electronic reporting is difficult if local modifications require initial and ongoing custom interfaces. An HIO could assist with transmission and standardizing message formats.

2. Nonmandated reporting of laboratory data: Not all diseases that might have public health significance are required to be reported, for example, influenza and gastrointestinal illnesses. HIOs could be set up to cooperate with the appropriate health authorities to support improved illness and prevention efficacy tracking.
3. Mandated reporting of physician-based diagnoses: As stated above, physician reporting is separate from laboratory and hospital reporting. Further, compliance ranges anywhere from 9 percent to 99 percent depending on the disease, provider training, and provider perception of the benefits. Effective use of HIE could ease the reporting burden on physicians while greatly increasing its reliability and validity.
4. Nonmandatory reporting of clinical data: HIO could assist with syndromic surveillance monitoring of nonreportable, nondiagnostic data such as the chief complaints from emergency departments or sales of over-the-counter medicines. Combined with reportable data and analyzed using the latest software, public health authorities could detect and address outbreaks earlier and more effectively.
5. Public health investigation: Public health authorities often become aware of cases that require investigation. With the availability of HIE, these authorities could use their legislated authority to perform these investigations.

6. Clinical care in public health clinics: In some jurisdictions public health authorities are required to provide healthcare. Having access to an HIE would enable these providers to access patient data needed for clinical care purposes.
7. Population-level quality monitoring: A regional HIO that included all of the providers of a region (not an ACO) could measure the quality of care delivered to members of the community, as well as the relevant outcomes.
8. Mass-casualty events: If the HIE were designed to receive ongoing admission, discharge, and transfer messages from registration systems, local disaster authorities could be empowered to allow designated centers to query information via HIE in response to requests from families and loved ones of persons who are missing.
9. Disaster medical response: The VHA was one of a handful of large healthcare providers with internal HIE that was utilized to make their patients' data available soon after Hurricane Katrina destroyed the physical infrastructure of New Orleans. Obviously, exchanging the data and having it exist in multiple locations results in redundancies that can be very beneficial in an emergency.
10. Public health alerting at the patient level: The use of HIE could make it easier for the public health authorities to track and ensure patients with designated conditions, such as tuberculosis or antibiotic-resistant organism infections, are tracked and treated appropriately when presenting for healthcare by new providers.
11. Public health alerting at the population level: HIOs could assist with provider alerts related to outbreak trends or necessary preventive services. If patients were also allowed access to a health portal powered by HIE, they could receive the same alerts and information for their given areas. This might help control the outbreak, spread, or prevention of selected diseases and conditions. (Shapiro et al. 2011)

The following researchers have documented the real benefits associated with HIOs. One set of researchers determined that information exchange facilitated by HIE between public health and clinical practice could improve decision-making and public health for infants with pertussis (Fine et al. 2010). Researchers in Louisiana sought to address the incidence of HIV/AIDS by using a bidirectional public HIE to facilitate the identification of patients with HIV/AIDS who had been lost to follow-up. When the patients reported to any ambulatory or inpatient facility, the clinicians were alerted to attempt to reengage the patients in clinical care. It was successful for 82 percent of the patients that have been lost to follow-up (Herwehe et al. 2012).

The studies just reviewed demonstrate that the theory of HIE use in public health can become a reality. With regulations and careful controls to prevent abuses of patient-level information, the use of HIE holds the promise of improved health and higher quality care for all citizens.

Case Study: Information Interoperability

A 50-year-old man with a history of substance abuse suffered a gunshot wound resulting in paraplegia. He presented in the emergency department (ED) at Hospital B in the evening with pain and fever after having been previously seen at four hospitals, including Hospital A where he left against medical advice. Medical records obtained from Hospital A by the ED staff at Hospital B where he had gone for treatment indicated that he previously had antibiotic treatment for sepsis.

Subsequently, the ED staff diagnosed the patient with sepsis resulting from

decubitus ulcers, he was started on an IV, given vancomycin and piperacillin and admitted to Hospital B. The next day he became unresponsive, underwent a code blue, and was transferred to ICU in a coma. The diagnosis was massive rupture of red blood cells due to a reaction to the antibiotics, specifically vancomycin and piperacillin. The result was brain injury and an unresponsive state necessitating discharge to a long-term care facility. The ICU eventually received records from the other hospitals that documented a previous cardiac arrest cause by an allergic reaction to the same antibiotics. When the ED staff asked the patient about allergic reactions, he reported that he had a blood transfusion reaction. However, this was transcribed as a statement indicating the patient had no allergies. (Adapted from Reider 2015)

Consider now these questions:

What would better system interoperability between hospital A and B change in this situation?

Can the ED staff improve the situation without interoperability? How?

Data need to follow an individual across continuity of care to ensure that the most appropriate care can be given in the appropriate place and time. In health and healthcare, HIEs are being used as a platform to allow health information to flow from one healthcare entity to another, which would benefit patients in variety of care settings by ensuring the most comprehensive and up-to-date data are available to assist healthcare providers in making decisions related to the patient's care. These very data can also be aggregated to guide policies and coordination of care for both specific communities and public. HIEs are available at many levels, from directly connecting healthcare entities within the same operational regions up to the federal level. Therefore, it is important that each healthcare entity adopts a standardized coding system to ensure that data exchanged are easily understood, accurate, and usable in care.

Chapter 3 Review Exercises

Answer the following questions.

1. Which version of ICD was published on January 1st, 2022?
   a. ICD version 9
   b. ICD version 10
   c. ICD version 11
   d. ICD version 12
   e. None of the above
2. Which of the following is not a potential public health benefit of HIE?
   a. Financial payment reform
   b. Disaster medical response
   c. Public health alerting
   d. Nonmandated reporting of laboratory data

Match the population or public health use of data to the correct example.

a. The Accountable Care Organization monitors whether smokers are receiving counseling.
b. The system ensures patients receive critical medications after a tornado.

c. Confirmed cases of HIV must be reported to the Department of Health.
d. The schools are being notified that there is an influenza outbreak in the state.

a. ___Mandatory laboratory result reporting
b. ___Population health monitoring
c. ___Disaster medical response
4. ___Public health alerts

Discussion.

7. Describe the history of health information exchange (HIE) in the United States.
8. List and define the different types of HIE/HIO structures.
9. Explain the advantages and disadvantages of the different types of HIE consent models.
10. Analyze the issues and challenges encountered when securing provider and consumer acceptance of HIE.
11. Describe the difference between population and public health.

REFERENCES

Adler-Milstein, J., A. Garg, W. Zhao, and V. Patel. 2021. A survey of health information exchange organizations in advance of a nationwide connectivity framework. ***Health Affairs*** 40(5):736–744. https://doi.org/10.1377/hlthaff.2020.01497.

Adler-Milstein, J., S. C. Lin, and A. K. Jha. 2016. The number of health information exchange efforts is declining, leaving the viability of broad clinical data exchange uncertain. ***Health Affairs*** 35(7):1278–1285. doi:10.1377/hlthaff.2015.1439.

American Health Information Management Association (AHIMA). 2017. ***Pocket Glossary for Health Information Management and Technology***. 5th ed. AHIMA Press.

Amatayakul, M. 2016. ***Health IT and EHRs.*** Chicago: AHIMA Press.

American Recovery and Reinvestment Act of 2009. Public Law 111-115.

Apathy, N. C., J. R. Vest, J. Adler-Milstein, J. Blackburn, B. E. Dixon, and C. A. Harle. 2021. Practice and market factors associated with provider volume of health information exchange. ***Journal of the American Medical Informatics Association*** 28(7):1451–1460. https://doi.org/10.1093/jamia/ocab024.

Banda, J. M., L. Evans, R. S. Vanguri, N. P. Tatonetti, P. B. Ryan, and N. H. Shah. 2016. Data descriptor: A curated and standardized adverse drug event resource to accelerate drug safety research. ***Science Data*** 3. doi:10.1186/s13326-016-0048-2.

Beckjord, E., R. Rechis, S. Nutt, L. Shulman, and B. Hesse. 2011. What do people affected by cancer think about electronic health information exchange? Results from the 2010 LIVESTRONG Electronic Health Information Exchange Survey and the 2008 Health Information National Trends Survey. ***Journal of Oncology Practice*** 7(4):237–241. doi:10.1200/JOP.2011.000324.

Blobel, B., W. Goossen, and M. Brochhausen. 2014. Clinical modeling—A critical analysis. ***International Journal of Medical Informatics*** 83(1):57–69. doi: 10.1016/j.ijmedinf.2013.09.003.

Blobel, B., F. Oemig, P. Ruotsalainen, and D. M. Lopez. 2022. Transformation of health and social care systems—An interdisciplinary approach toward a foundational architecture. ***Frontiers in Medicine*** 9:802487. doi:10.3389/fmed.2022.802487.

Bowman, S., R. Butler, K. Giannangelo, N. Leon-Chisen, and R. Schichilone. 2016. Putting the ICD-10-CM/PCS GEMs into practice (updated). ***Journal of AHIMA*** 87(1):48–53.

Centers for Disease Control and Prevention (CDC). 2015. "Ebola (Ebola virus)." https://www.cdc.gov/vhf/ebola/pdf/ebola-factsheet.pdf.

Centers for Disease Control and Prevention (CDC). 2022a. "COVID-19 timeline." https://www.cdc.gov/museum/timeline/covid19.html.

Centers for Disease Control and Prevention (CDC). 2022b. "National Notifiable Diseases Surveillance System." https://www.cdc.gov/nndss/data-statistics/index.html.

Centers for Medicare and Medicaid Services (CMS). 2012. Electronic Health Record Incentive Program—Stage 2; Health Information Technology: Standards, Implementation Specifications, and Certification Criteria for Electronic Health Record Technology, 2014 ed.; Revisions to the Permanent Certification Program for Health Information Technology; Final Rules. ***Code of Federal Regulations.*** http://www.gpo.gov/fdsys/pkg/FR-2012-09-04/pdf/2012-21050.pdf.

Centers for Medicare and Medicaid Services (CMS). 2015a. "CMS fact sheet: EHR incentive programs in 2015 and beyond." https://www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2015-Fact-sheets-items/2015-10-06-2.html.

Centers for Medicare and Medicaid Services (CMS). 2015b. "EHR incentive programs: 2015 to 2017 (modified Stage 2) overview." https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Downloads/2015_EHR2015_2017.pdf.

Centers for Medicare and Medicaid Services (CMS). 2015c. "EHR incentive programs in 2015 through 2017: Health information exchange." https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Downloads/2016_HealthInformationExchange.pdf.

Centers for Medicare and Medicaid Services (CMS). 2015d. "***Data segmentation.***" https://www.healthit.gov/providers-professionals/data-segmentation-overview.

Centers for Medicare and Medicaid Services (CMS). 2015e. "***Enabling privacy: Data segmentation.***" https://www.healthit.gov/providers-professionals/ds4p-initiative.

Centers for Medicare and Medicaid Services (CMS). 2022. "Promoting interoperability programs." https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms.

Dhopeshwarkar, R. V., L. M. Kern, H. C. O'Donnell, A. M. Edwards, and R. Kaushal. 2012. Health care consumers' preferences around health information exchange. ***Annals of Family Medicine*** 10(5):428–434.

Dimitropoulos, L., V. Patel, S. Scheffler, and S. Posnack. 2011. Public attitudes toward health information exchange: Perceived benefits and concerns. ***American Journal of Managed Care*** 17:SP111–116.

Dixon, B. E., S. J. Grannis, C. McAndrews, A. A. Broyles, W. Mikels-Carrasco, A. Wiensch, J. L. Williams, U. Tachinardi, and P. J. Embi. 2021. Leveraging data visualization and a statewide health information exchange to support COVID-19 surveillance and response: Application of public health informatics, ***Journal of the***

American Medical Informatics Association 28(7):1363–1373. https://doi.org/10.1093/jamia/ocab004.

Federal Register. 2018. "Medicare Program; Hospital inpatient prospective payment systems for acute care hospitals and the long-term care hospital prospective payment system and policy changes and Fiscal Year 2019 rates; Quality reporting requirements for sspecific providers; Medicare and Medicaid Electronic Health Record (EHR) incentive programs (promoting interoperability programs) requirements for eligible hospitals, critical access hospitals, and eligible professionals; Medicare cost reporting requirements; and physician certification and recertification of claims." https://www.federalregister.gov/citation/83-FR-41635.

Federal Register. 2019. "Medicare and Medicaid Programs; Patient Protection and Affordable Care Act; Interoperability and patient access for Medicare Advantage organization and Medicaid managed care plans, state Medicaid agencies, CHIP agencies and CHIP managed care entities, issuers of qualified health plans in the federally-facilitated exchanges and health care providers." https://www.federalregister.gov/documents/2019/03/04/2019-02200/medicare-and-medicaid-programs-patient-protection-and-affordable-care-act-interoperability-and.

Fine, A. M., B. Y. Reis, L. E. Nigrovic, D. A. Goldmann, T. N. LaPorte, K. L. Olson, and K. D. Mandl. 2010. Use of population health data to refine diagnostic decision-making for pertussis. ***Journal of the American Medical Informatics Association*** 17(1):85–90. doi:10.1197/jamia.M3061.

Friedman, D. J., and R. G. Parrish. 2010. The population health record: Concepts, definition, design, and implementation. ***Journal of the American Medical Informatics Association*** 17(4):359–366. doi:10.1136/jamia.2009.001578.

Frisse, M. E., K. B. Johnson, H. Nian, C. L. Davison, C. S. Gadd, K. M. Unertl, P. A. Turri, and Q. Chen. 2012. The financial impact of health information exchange on emergency department care. ***Journal of the American Medical Informatics Association*** 19(3):328–333. doi:10.1136/amiajnl-2011-000394.

Gadd, C. S., Y. X. Ho, C. M. Cala, D. Blakemore, Q. Chen, M. E. Frisse, and K. B. Johnson. 2011. User perspectives on the usability of a regional health information exchange. ***Journal of the American Medical Informatics Association*** 18(5):711–716. doi:10.1136/amiajnl-2011-000281.

Giannangelo, K. 2015. ***Healthcare Code Sets, Clinical Terminologies, and Classification Systems***. 3rd ed. Chicago: AHIMA.

Gourevitch, M. N., T. Cannell, J. I. Boufford, and C. Summers. 2012. The challenge of attribution: Responsibility for population health in the context of accountable care. ***American Journal of Public Health*** 102(Suppl. 3):S322–S324. doi:10.2105/AJPH.2011.300642.

Gray, P. 2011. ***Consent Options for HIE in Texas.*** Houston: Texas Health Services Authority. http://hietexas.org/resources/policy-guidance.

HL7.org. 2015. "***FHIR overview.***" https://www.hl7.org/fhir/overview.html.

Haas, S. A. 2016. Developing staffing models to support population health management and quality outcomes in ambulatory care settings. ***Nursing Economics*** 34(3):126–133.

Hardcastle, L., K. Record, P. Jacobson, and L. Gostin. 2011. Improving the population's health: The Affordable Care Act and the importance of integration. ***Journal of Law,***

Medicine and Ethics 39(3):317–327. doi:10.1111/j.1748-720X.2011.00602.x.
HcPro. 2016. Providers preparing for MIPS. ***Doctor's Office*** 35(7):1–5.
Healtheway, Inc. 2012. http://healthewayinc.org/.
HealthIT. 2020. "HHS announces funding for health information exchanges to support public health agencies—STAR HIE program funds five organizations to improve interoperability of health data." https://www.hhs.gov/about/news/2020/09/30/hhs-announces-funding-for-health-information-exchanges-to-support-public-health-agencies.html.
Health Insurance Portability and Accountability Act of 1996. Public Law 104-191.
Herwehe, J., W. Wilbright, A. Abrams, S. Bergson, J. Foxhood, M. Kaiser, L. Smith, K. Xiao, A. Zapata, and M. Magnus. 2012. Implementation of an innovative, integrated electronic medical record (EMR) and public health information exchange for HIV/AIDS. ***Journal of the American Medical Informatics Association*** 19(3):448–452.
Hess, J. 2011. The brave, new world of HIEs. ***Healthcare Financial Management*** 65(2):44–48.
Hincapie, A. L., T. L. Warholak, A. C. Murcko, M. Slack, and D. C. Malone. 2011. Physicians' opinions of a health information exchange. ***Journal of the American Medical Informatics Association*** 18(1):60–65. doi:10.1136/jamia.2010.006502.
ICare. 2016. Integrated care collaboration. http://icc-centex.org/.
International Health Terminology Standards Development Organization. 2013. "SNOMED-CT user guide." http://ihtsdo.org/fileadmin/user_upload/doc/.
Joshi, J. 2011. Clinical value-add for health information exchange (HIE). ***Internet Journal of Medical Informatics*** 6(1):1.
Khurshid, A., M. L. Diana, and R. Jain. 2015. Health information exchange readiness for demonstrating return on investment and quality of care. ***Perspectives in Health Information Management*** 12(Fall): 1–15.
Kim, J., R. L. Ohsfeldt, L. D. Gamm, T. A. Radcliff, and L. Jiane. 2016. Hospital characteristics are associated with readiness. ***The Journal of Rural Health*** 9:1–9. doi:10.1111/jrh.12193.
Kyounghae, K., J. S. Choi, C. Eunsuk, C. L. Nieman, J. Jin Hui, F. R. Lin, L. N. Gitlin, and H. Hae-Ra. 2016. Effects ofcommunity-based health worker interventions to improve chronic disease management and care among vulnerable populations: A systematic review. ***American Journal of Public Health*** 106(4):e3–e28. doi:10.2105/AJPH.2015.302987.
LaMonica, H M., T. A. Davenport, A. E. Roberts, and I. B. Hickie. 2021. "Understanding technology preferences and requirements for health information technologies designed to improve and maintain the mental health and well-being of older adults: participatory design study." JMIR aging 4, no. 1 (2021): e21461.
McCann, E. 2014. Landmark year ahead for HIE. ***Healthcare IT News.*** http://www.healthcareitnews.com/news/landmark-year-ahead-hie.
Moreno-Conde, A., F. Jódar-Sánchez, and D. Kalra. 2015. Requirements for clinical information modelling tools. ***International Journal of Medical Informatics*** 84(7):524–536. doi:10.1016/j.ijmedinf.2015.03.005.
Nandigam, H. 2016. "Mapping Systematized Nomenclature of Medicine-Clinical Terms (SNOMED CT) to International Classification of Diseases, Tenth Revision, Clinical

Modification (ICD-10-CM): Lessons Learned from Applying the National Library of Medicine's Mappings." Perspectives in Health Information Management (2016): 1

National Institute of Standards and Technology (NIST). 2016. "Standards and testing FAQs." http://healthcare.nist.gov/use_testing/faq.html.

National Library of Medicine, National Institutes of Health (NLM). 2016. "***SNOMED CT to ICD-10-CM.***" https://www.nlm.nih.gov/research/umls/mapping_projects/snomedct_to_icd10cm.html.

Oachs, P., and A. Watters. 2020. ***Health Information Management: Concepts, Principles, and Practice***, 6th ed., edited by P. Oachs and A. Watters. Chicago: AHIMA.

Office of the National Coordinator, DHHS (ONC). 2012. "Consumer engagement in health information exchange." https://www.healthit.gov/sites/default/files/consumer_mediated_exchange.pdf.

Office of the National Coordinator, DHHS (ONC). 2013a. "Vocabulary Task Force." http://www.healthit.gov/policy-researchers-implementers/vocabulary-task-force.

Office of the National Coordinator, DHHS (ONC). 2013b. "HealthIT.hhs.gov: State Health Information Exchange Program." https://www.healthit.gov/policy-researchers-implementers/state-health-information-exchange.

Office of the National Coordinator, DHHS (ONC). 2016. "HealthIT.hhs.gov: HIE bright spots—Lab Exchange." https://www.healthit.gov/sites/default/files/lab_exchange_bright_spots_synthesis_final_09302013.pdf.

Office of the National Coordinator for Health Information Technology (ONC). n.d. Health Information Technology Advisory Committee. https://www.healthit.gov/hitac/committees/health-information-technology-advisory-committee-hitac.

Patel, V., R. Dhopeshwarkar, A. Edwards, Y. Barrón, J. Sparenborg, and R. Kaushal. 2012. Consumer support for health information exchange and personal health records: A regional health information organization survey. ***Journal of Medical Systems*** 36(3):1043–1052. doi:10.1007/s10916-010-9566-0.

Pevnick, J., M. Claver, A. Dobalian, S. Asch, H. Stutman, A. Tomines, and P. Fu. 2012. Provider stakeholders' perceived benefit from a nascent health information exchange: A qualitative analysis. ***Journal of Medical Systems*** 36(2):601–613. doi:10.1007/s10916-010-9524-x.

Rahurkar, S., J. R. Vest, J. T. Finnell, and B. E. Dixon. 2021. Trends in user-initiated health information exchange in the inpatient, outpatient, and emergency settings, ***Journal of the American Medical Informatics Association*** 28(3): 622–627. https://doi.org/10.1093/jamia/ocaa226.

Reider, J. 2015. "Agency for healthcare research and quality. The risks of absent interoperability: Medication-induced hemolysis in a patient with a known allergy." https://psnet.ahrq.gov/webmm/case/359/the-risks-of-absent-interoperability-medication-induced-hemolysis-in-a-patient-with-a-known-allergy.

Shapiro, J., F. Mostashari, G. Hripcsak, N. Soulakis, and G. Kuperman. 2011. Using health information exchange to improve public health. ***American Journal of Public Health*** 101(4):616–623. doi:10.2105/AJPH.2008.158980.

Tompkins, D. K. 2016. ICD-10 opens documentation gaps. ***Journal of AHIMA*** 87(3):54–55.

Tsai, J., and R. Rosenheck. 2012. Use of the internet and an online personal health record system by US veterans: Comparison of Veterans Affairs mental health service users and other veterans nationally. ***Journal of the American Medical Informatics Association*** 19(6):1089–1094. doi:10.1136/amiajnl-2012-000971.

US Department of Health and Human Services (HHS). 2010. Health information technology: Initial set of standards, implementation specification and certification criteria for electronic health record technology. ***Code of Federal Regulations.*** http://edocket.access.gpo.gov/2010/pdf/2010-17210.pdf.

US Department of Health and Human Services, Centers for Medicare and Medicaid Services, Medicare Program (HHS CMS). 2011. "Medicare shared savings program: Accountable Care Organizations, final rule." http://www.gpo.gov/fdsys/pkg/FR-2011-11-02/pdf/2011-27461.pdf.

Vest, J. R., and L. D. Gamm. 2010. Health information exchange: Persistent challenges and new strategies. ***Journal of the American Medical Informatics Association*** 17(3):288–294. doi:10.1136/jamia.2010.003673.

Vest, J., L. Gamm, R. Ohsfeldt, H. Zhao, and J. Jasperson. 2012. Factors associated with health information exchange system usage in a safety-net ambulatory care clinic setting. ***Journal of Medical Systems*** 36(4):2455–2461. doi:10.1007/s10916-011-9712-3.

Vest, J., and J. Jasperson. 2012. How are health professionals using health information exchange systems? Measuring usage for evaluation and system improvement. ***Journal of Medical Systems*** 36(5):3195–3204. doi:10.1007/s10916-011-9810-2.

Vest, J. R., and B. A. Kash. 2016. Differing strategies to meet information-sharing needs: publicly supported community health information exchanges versus health systems' enterprise health information exchanges. ***Milbank Q.*** 941:77–108.

World Health Organization (WHO). 2013a. ***International Classification of Diseases.*** http://www.who.int/classifications/icd/en/.

World Health Organization (WHO). 2013b. ***International Classification of Diseases, 11th revision.*** www.who.int/classifications/icd11/.

World Health Organization (WHO). 2022. Classifications. http://www.who.int/classifications/icd/en/.

Yeager, D. 2016. The challenges of population health management. ***For the Record*** 28(2):22–25.

ADDITIONAL RESOURCES

American Health Information Management Association (AHIMA). 2016a. "***Three keys to simplifying interoperability in healthcare.***" http://journal.ahima.org/2016/04/15/three-keys-to-simplifying-interoperability-in-healthcare-sponsored/.

American Health Information Management Association (AHIMA). 2016b. Survey: ICD-10-CM/PCS has reduced coding productivity. ***Journal of AHIMA*** 87(1):9.

American Medical Association (AMA). 2013. "Current procedural terminology." https://www.ama-assn.org/about-us/current-procedural-terminology-cpt-editorial-panel.

Averill, R. F., R. L. Mullin, B. A. Steinbeck, N. L. Goldfield, M. T. Grant, and R. R. Butler. 2016. ***Development of the ICD-10 Procedure Coding System (ICD-10-***

PCS).

Bielby, J. A. 2016. Continuing on the intriguing ICD-10 journey. ***Journal of AHIMA*** 87(5):50–52.

Bowman, S. 2008. Why ICD-10 is worth the trouble. ***Journal of AHIMA*** 79(3):24–29.

Centers for Disease Control and Prevention (CDC). 2013a. "***ICD—ICD-9-CM—Coordination and Maintenance Committee."*** http://www.cdc.gov/nchs/icd/icd9cm_maintenance.htm.

Centers for Disease Control and Prevention (CDC). 2013b. International Classification of Diseases, 10th rev, Clinical Modification (ICD-10-CM). http://www.cdc.gov/nchs/icd/icd10cm.htm.

Centers for Disease Control and Prevention (CDC). 2016. "***Mortality data."*** http://www.cdc.gov/nchs/nvss/deaths.htm.

Centers for Medical and Medicaid Services (CMS). 2016. "***ICD-10-PCS***." https://www.cms.gov/Medicare/Coding/ICD10/downloads/pcs_final_report2012.pdf.

Council of State and Territorial Epidemiologists. 2012. "***List of nationally notifiable conditions.***" www.cste.org/resource/resmgr/PDFs/CSTENotifiableConditionListA.pdf.

Department of Health and Human Services (HHS). 2012. Administrative simplification: Adoption of a standard for a unique health plan identifier; Addition to the national provider identifier requirements; and a change to the compliance date for the International Classification of Diseases, 10th ed. Final rule. ***Federal Register.*** https://www.federalregister.gov/documents/2012/09/05/2012-21238/administrative-simplification-adoption-of-a-standard-for-a-unique-health-plan-identifier-addition-to.

Frisse, M. E. 2010. Health information exchange in Memphis: Impact on the physician-patient relationship. ***Journal of Law, Medicine and Ethics*** 38(1):50–57. doi:10.1111/j.1748-720X.2010.00465.x.

Logical Observation Identifiers Names and Codes (LOINC®). 2016. http://loinc.org/.

Mayo Clinic Informatics: CIMI. 2013. http://informatics.mayo.edu/CIMI/index.php/Main_Page.

National Library of Medicine, National Institutes of Health (NLM). 2016. "RxNorm." https://www.nlm.nih.gov/research/umls/rxnorm/.

Ramanathan, T., C. Schmidt, A. Menon, D. Pepin, and M. Penn. 2015. "Issue Brief: Federal public health laws supporting data use and sharing." https://www.cdc.gov/phlp/docs/datasharing-laws.pdf.

Chapter

4

Implementing Healthcare Information Systems

By Suzanne Paone, DHA, MBA, RHIA

Learning Objectives

- Apply principles of people, process, and technology to health systems implementation
- Relate systems adoption to effective implementation strategies
- Articulate the importance of leadership engagement
- Analyze the systems development life cycle as a process framework for implementation
- Contrast various types of technology implementation initiatives
- Analyze the value of effective health systems implementation

KEY TERMS

Adaptable solutions
Adoption
Analytics and reporting verification
Capital costs
Champion
Conversion
Data conversion
Data level security
Documentation
Downtime and disaster planning verification
Implementation
Operational costs
Organizational governance
Post-conversion support

Program
Program manager
Project
Project manager
Software solution configuration
Sponsor
Stakeholder
Superuser
Sustainment
System building and development
Systems development life cycle (SDLC)
Testing
Training

The health sector of the economy is comprised of organizations with a common aim of improving the delivery of health products and services. To support transactional business operations, analytics demands, and transformational strategies of organizations such as health providers, insurance companies, private practitioners, consultancy organizations, and industry development organizations, health-related clinical and business systems need to be utilized effectively. The American Medical Informatics Association defines informatics as the use of data information and knowledge to improve the delivery of health in society promote transformation (AMIA 2022) Recall that health informatics (HI) involves the use of specific technology platforms and data such as clinical information systems (CIS) and electronic health record (EHR) systems to create data supporting organizational functions. The ability for organizations and societies to utilize health data appropriately depends on the implementation of a variety of technology solutions within a framework of people, process, and technology effectiveness.

People and Health Systems Implementation

The work involved in introducing technologies of all types into health organizations involves a myriad of different people. Some individuals are in unique roles either within the organization or external to the organization and possess a high level of influence or interest regarding the success of technology-related initiatives. In addition, groups of individuals that perform similar job functions or roles within an organization have a highly vested interest in the efficient implementation of technology solutions such as software systems, daily use hardware, secure network infrastructure, and mobile technologies. Despite advanced levels of automation, the proficiency of individuals and groups of individuals related to technology directly influences the ability of organizations to manage costs, deliver quality products and services, and be competitive.

Stakeholders

An individual, formal role such as chief nursing officer or groups of individuals or entities may be **stakeholders** in the implementation of health system solutions. Stakeholders have a vested interest in both the implementation of a technology solution and the capability for a solution to provide valuable data to meet the objectives of the organization. Individuals or groups of individuals serving as stakeholders can be internal or external to the organization in which the health systems solutions are being implemented (Georgiou et al. 2018). Stakeholders' level of interest and influence in any

implementation process varies based on the background and experience of stakeholders, the organizational culture and goals, and imperatives of the implementation (Kettlewell et al. 2018).

A stakeholder's level of interest in the implementation and level of influence in the organization are two parameters by which stakeholder positioning is measured (Andrews et al. 2022). For example, the division head of the noninvasive cardiology department in a multisite hospital might have a high level of interest in the implementation of a new software documentation and orders system in the emergency departments of the hospitals in the network. Although this stakeholder's interest is high, the level of influence that this stakeholder has on the implementation of the emergency department software is low. What this means to the leaders sponsoring the implementation and the implementation team is that the stakeholder should be on the communication list of those informed about the progress of the implementation but not necessarily included in direct frequent communications, updates, or decision-making committees.

page 87

Implementations of complex transactional business software solutions such as EHRs, personal health records (PHRs), human resource management systems, supply chain management systems, accounting and financial systems, or telemedicine applications involve specialized types of stakeholders. The **sponsors** of these implementations are the individuals or groups of individuals who provide financial resources that enable successful and sustainable implementations of health system solutions. A **champion** is a specialized stakeholder who has a defined role to provide one-on-one support to influential groups of stakeholders. For example, a physician champion in the surgery department in an academic hospital setting facilitates the continued training and day-to-day support of new residents using surgery system documentation templates in the new EHR system.

Types of Stakeholders

- Internal stakeholders; staff associates, providers, management team, executives, board of directors, advisory boards
- External stakeholders; customers, vendors, investors, governments, communities
- Sponsors
- Champions

Implementation Teams

Implementation involves the process of delivering technology, data sets, and processes to a business function. Technologies such as EHRs are implemented through a series of steps for use in various settings of clinical care; platforms such as human resources systems are implemented to support administrative functions such as recruiting and payroll in the same settings of care. Most often, technology-based solutions of all types are implemented in various health-related settings using team structures. Many implementation initiatives involve adding new value to an organization in the framework of organizational strategies. When work is defined in this way, often the implementation is taking place in the framework of a project. A **project** begins and ends with a distinct set of goals and objectives and a defined scope of work. In larger, more complex settings, multiple projects that have relationships to each other might involve the implementation of multiple types of technology solutions simultaneously. This group of projects with interdependencies and relationships between project work and goals is a **program**. Whether the implementation happens in a framework of a single project, two or three projects, or a formal program, most often people operate within team structures that are directed by leadership groups, committees, or individuals.

page 88

Implementation Structures

Organizational governance involves people or groups of people in specialized committees or day-to-day operating committees to make decisions within organizations. These decisions might be strategic in nature regarding the overall direction of the organization in the marketplace or they might be operational in nature. For example, a medical executive committee is a requirement of an accredited health system. This group of individuals has decision-making and policy setting responsibility for many aspects of practicing medicine. If an upgrade to the digital health record platform is being implemented, this committee might make decisions regarding policy, such as health specialists or primary care physicians using the problem list functionality within the EHR software. In a smaller organizational setting such as an independent endocrinology office, governance responsibility for this type of decision-making during the implementation of an EHR might lie with the practice's medical director and administrator. Either way, the decisions needed for implementation teams to move forward are rendered by governance structures.

Implementation Team Roles

When a health system technology-based solution is being implemented, depending on the scope and size of the implementation initiative (single projects versus multiple projects or programs), several teams are formed to do the work involved in completing the implementation. Teams could include technical staff who are responsible for the requirements design and implementation of technical system interfaces between different types of software platforms. Another team might include business analysts and end users from the supply chain organization of a hospital system who are involved in the upgrade of the enterprise supply chain software that runs clinics, hospitals, and physician practices. Teams of physicians might work on medical content needed for a telemedicine platform to be implemented across the service region of a community hospital. When working on an implementation team, individuals have specific responsibilities to complete work on time and within established budget parameters, within their areas of subject matter expertise. Regardless of the nature of the team structure, many sizable implementations designate individuals in leadership roles at the team level.

A group of teams implementing solutions in the framework of project or projects will likely have a **project manager**. Typically, the project manager has responsibility for the entirety of the project including financial accountability, quality delivery, and timeliness of the implementation. If the implementation is sizable, several teams might be organized under the guidance of team leads. These team leads take responsibility for the work of teams, ensuring that work is completed on time, within budget constraints, and within established quality parameters. If multiple projects constitute a formal implementation program, individuals may serve in roles called program managers. The **program manager** reports to organizational sponsors and organizational governance groups regarding implementation progress specific to the quality delivery, meeting the goals of the organization, and financial responsibility for the cost of the program.

Organizational Leaders

Implementation projects and programs are sponsored by organizations to facilitate the successful execution of organizational strategies. In the digital health economy, organizational strategies are dynamic based on changing environmental parameters that affect health organizations of all types. Strategic direction in health organizations is less an exclusive function of top-down organizational governance from the board of directors, but is also established by incorporating innovative ideas of leaders and staff associates in organizations (Halal et al. 2022). In this dynamic strategic environment, technology solutions involving software, hardware, internet-based modalities, predictive analytical

business intelligence, and other solutions provide the capability for health organizations to execute high-level strategies such as expanding market share, establishing low-cost leadership, differentiating in areas of innovation, and providing better customer service. Leaders in high-performing health organizations understand that strategies rely on the effective and sustainable use of technology-based systems solutions.

Levels of Strategy

Organizational leaders such as boards of directors, advisory boards, and executives develop strategies that allow organizations to be not only viable today in executing their mission, but increasingly viable as environmental factors change in markets in which organizations operate. Organizational strategies that must be facilitated by the implementation of efficient technology solutions and ongoing sustainment of those solutions include strategies such as growth into new markets, differentiation of products and services based on innovation and customer service, and partnership joint venture or merger activities. Regardless of the strategy used to execute the organization's mission, decision makers rely on timely and accurate information to execute strategic objectives (Abdekhoda et al. 2019). Appropriate organizational decision-making relies on implementing systems inclusive of software, hardware, networks, and data platforms that reliably produce data that can be transformed into useful information. Organizational strategies are supported by day-to-day business strategies, and day-to-day business strategies are supported by information technology strategies. The successful execution of a viable information technology strategy relies on the implementation of not only new technologies, but the continued updating and upgrading of appropriate technical solutions to support information needs.

Sponsors, Champions, and Decision-Making

Governance groups determine the goals of a project or program implementation. In collaboration with vendors, possibly consultants, and internal project and program staff, sponsors develop well-informed ideas of how technology implementation can benefit the organization. For example, the implementation of an integrated PHR in a healthcare delivery system has vast potential related to both administrative and clinical goals. The chief financial officer (CFO) of the organization might expect to see customer convenience functions deployed early in the project, such as online bill payment to reduce days in accounts receivable and eliminate costs associated with check processing for customer co-pays. Likewise, the chief medical officer (CMO) of the health system looks at the implementation of the PHR as an opportunity to provide physicians the capability to improve the health management of chronically ill patients such as people with congestive heart failure and diabetes by implementing integrated home monitoring devices connected through the PHR to the EHR platform. These leadership goals are not mutually exclusive but complementary. A well-organized implementation approach involving multiple projects allows the organization to save money by implementing bill payment online while also offering tethered medical devices to chronically ill patients.

One of the challenges with implementing technology solutions is ensuring that timely and well-informed decisions are made by appropriate governance groups and individuals to allow the implementation of technical solutions to move forward (see Table 4.1). Project managers, program managers, sponsors, and champions working within implementation teams are responsible for articulating the decisions that need to be made to governance groups. Project teams should not be shy about collating and articulating the critical policy-related decisions and or policy changes that need to be made for implementation activities to progress. For example, in the same implementation involving the integrated PHR, transformational capabilities exist regarding the release of laboratory and other types of test results directly to patients.

In this example, governance groups need to revisit current policies regarding the release of results to patients and consider patient preferences and appropriate operational policies to determine how automated results will be directly released to patients following the implementation of the PHR. The implementation team should provide decision makers with an idea of the systems capabilities as well as information regarding how similar organizations that have implemented PHRs have handled the release of test results with similar software. This information about software capabilities and organizational benchmarking facilitates evidence-based and thoughtful policy development as the health system transforms its provider–customer relationships.

Adoption and Implementation

Organizational goals are achieved, and industry transformation occurs when stakeholders effectively use technology platforms in a sustainable manner. **Adoption** means the sustainable day-to-day use of technologies that enable organizations to produce quality data that are transformed into information and knowledge-based actions. The appropriate adoption of technology-based solutions in the health industry has the capability to be vastly transformative. Professional organizations such as the American Health Information Management Association support standards, methodologies, and best practices to enable the individual and organizational adoption of technologies to enable transformation (AHIMA 2022). Efficient technology adoption is critical for all divisions and departments of a high-performing organization.

Table 4.1. Sample leadership-level stakeholder interests

Role	Type of health organization	Example health system implementation	Stakeholder interest
Chief executive officer (CEO)	Integrated delivery health network	New EHR platform	High interest and high influence; clinical quality, client flow in care settings, effect on the revenue cycle such as accurate coding and claims to insurance companies
Chairman, board of directors	Hospital system	New enterprise resource planning (ERP) supply chain platform	Low interest and high influence; general expectation that this platform would increase industry cost competitiveness
Chief human resource officer (CHRO)	Academic medical center	Upgrade to the existing human resource information system (HRIS) software	Low interest and high influence; concerns about effect on daily operations of the human resource organization as well as define strategic objectives for human resources to be more efficient and competitive
Advisory board director	Independent multipractice physician office	New practice digital health record	Low interest and high influence; expects that the implementation of a new digital health record will align with the strategy of the physician practice and enhance their competitive position
Chief nursing officer	Community health system	New clinical staffing system	High interest and high influence; expects that new staffing software allowing self-scheduling preferences will help with competitive workforce concerns

Chief marketing officer	Health insurance company	New web content management system platform	High interest and high influence; the contact management system provides the ability for the organization to efficiently decentralize contact management and reduce costs associated with the organization's web presence
Medical director	Multisite chiropractor offices	New client portal	High interest and high influence; increased competitiveness based on client ability to schedule appointments, message clinicians, and receive customized content

page 92

We know from studies conducted over decades that one of the most prevalent models for technology adoption is the Rogers Adoption Curve. Over the course of decades of implementation of systems and for many industry sectors including the health sector, this theory stands the test of time in terms of explaining the way in which groups of persons efficiently adopt new technology innovation (Soh and Grover 2020). This model divides stakeholders in an organization where technology is being implemented into three general groups: early adopters, the majority, and laggards or those who resist technology adoption. Stakeholders accept new technology and redefined processes based on their tendency to fall into one of these categories. Early adopters readily accept new solutions and sustain their use in operations (Kalayou et al. 2020). End users who form the majority follow new processes established by early adopters and allow organizations to add new functions into existing workflows (Bangash et al. 2020). Laggards are those who resist technology-based change and can prevent the effective progression of systems such as EHRs in an organization (Rumball-Smith et al. 2020). Early adopters and laggards comprise approximately 35 percent of the totality of people adopting the implementation of a new technology; approximately 65 percent of people fall into the middle of a normal distribution as majority adopters. During the planning stage of the systems development life cycle, it is critical to formulate specific strategies to engage individuals who are early adopters and laggards. Implementation teams do this in tandem with project and program sponsors to facilitate organizational adoption during the implementation phase.

page 93

Ensuring a comprehensive stakeholder strategy is not the only way to facilitate effective adoption of technology during a technology implementation and after the implementation is complete. As discussed in one recent study (Janssen et al. 2021), there are many examples in the literature of health organizations who have positive people-based adoption but encounter other issues that serve as barriers to an effective implementation. The lack of a sustainable technology support and maintenance framework is a major barrier to adoption in many health organizations, especially organizations with limited technology resourcing or in geographic areas where staffing technology talent is a challenge. Best practices in the post-implementation sustainment phase that occurs after the technology goes live are addressed below (Tsai et al. 2020).

Promoting Efficient Adoption

- Implementation team leaders communicate efficiently; stakeholder needs are analyzed for frequency and type of communication they want to receive
- Prior to implementation during the planning stage, sponsors and implementation leaders formulate strategies for engaging both early adopters and laggards
- Implementation planning leaders formulate extensive sustainability plans post go live; focus on ensuring efficient support once the technology system goes live

- Choose champions who are not only early adopters, but also laggards with inherent resistance to the new technology

Check Your Understanding 4.1

Answer the following questions.

1. Which statement most accurately describes the role of the sponsor during the software implementation process?
   a. The subject matter expert who will provide guidance as to how to design the software
   b. An individual who provides financial resources to support the software implementation project
   c. The individual responsible for signing the contract with a software vendor company
   d. The committee managing the alignment of physician interests during the implementation process
2. Which phrase is the best definition of a project?
   a. It has a beginning and an end
   b. It has a specific goal
   c. Tt has a defined scope of work
   d. All of the above

3. Identify the scenario which should be submitted to a governance committee who oversees the implementation of a new HRIS (human resource management system) in a large academic health system
   a. Decisions regarding what days of the week supervisors will sign off on payroll hours for employees
   b. The design of the recruiting website used for open job positions advertised outside of the health system
   c. The policy that determines which level of digital executive signature is required to approve requests for new staff positions
   d. Designing the number of barcode readers in each department for secure staff entry into buildings
4. Identify the most appropriate phrase regarding the stakeholder interest of the Chief Nursing Officer during the implementation of a new electronic health record in a hospital.
   a. High interest, high influence
   b. Low interest, low influence
   c. High interest, low influence
   d. Low interest, high influence
5. The implementation team has implemented the new EHR effectively in three departments within an academic medical center (surgery, medicine and urology). Which term best describes the Chief of the Department of Ophthalmology who tells all of this fellow department chiefs that the new EHR is inadequate and should not be implemented.
   a. An early adopter
   b. A sponsor
   c. An end user
   d. A laggard

Processes in Health Systems Implementation

Revisiting the definition of HI, we realize that the appropriate utilization of data and information to create transformational knowledge in the health industry depends on people, process, and technology. Stakeholders, sponsors, and champions as well as persons performing implementation activities and sustaining health system solutions implement these solutions optimally in proven process frameworks. The effective planning and implementation and sustainment of these solutions depends on balancing the realities of time, cost, and expectations of the business in terms of addressing both clinical and administrative quality metrics. As evidenced in all industries, transformational organizational activities focused on increasing quality metrics to the highest levels are best obtained and sustained when technology aligns with optimal process (Land et al. 2012).

The Systems Development Life Cycle

The health sector can best support informatics goals and objectives by leveraging existing proven frameworks that provide support for transformational technology-based initiatives. Initiatives are best deployed in project- or program-based frameworks using proven project methodologies. These might include traditional project management methods or contemporary methods such as rapid application deployment techniques (Lai 2017; Unhelkar 2016). Regardless of the project or program methodology, technology-based solutions such as software solutions, device implementations, and analytics implementations are optimally delivered to the business and sustained using the construct of the **systems development life cycle** (SDLC); this concept is standards based and used across multiple sectors of the economy to optimize the implementation of technology and the sustained informatics goals of transforming data into information and knowledge-based intelligence (Radack 2022.).

Holistic Nature of the SDLC

The SDLC is made up of several repetitive phases involving standard activities utilized to plan for, implement, and support new products, services, and technologies. There are many versions of the SDLC and literature. Figure 4.1 outlines a summative version of the SDLC utilized to plan, implement, and sustain health system solutions in a myriad of health-related organizational settings. The SDLC approach is appropriate for any type of technology or data-related implementation. For example, a provider health system CEO and CFO begin discussions with the board of directors about increased organizational performance expectations regarding clinical quality performance, financial performance, and customer-facing product lines that need to be developed such as telemedicine. Realizing that current technology platforms utilized in the health system are not optimally supporting the increased competitive pressure on the health system, the board of directors and executive team allocate internal resources working in concert with an experienced consultant to review the market and understand the cost and benefits for the organization of pursuing the implementation of new EHR platforms. In this case, the SDLC commences with the planning phase. Leaders have a vested interest in the successful implementation of technology solutions such as EHRs based on the need to meet the expectations of the community and their customers (Hagland 2017; Hartley et al. 2020).

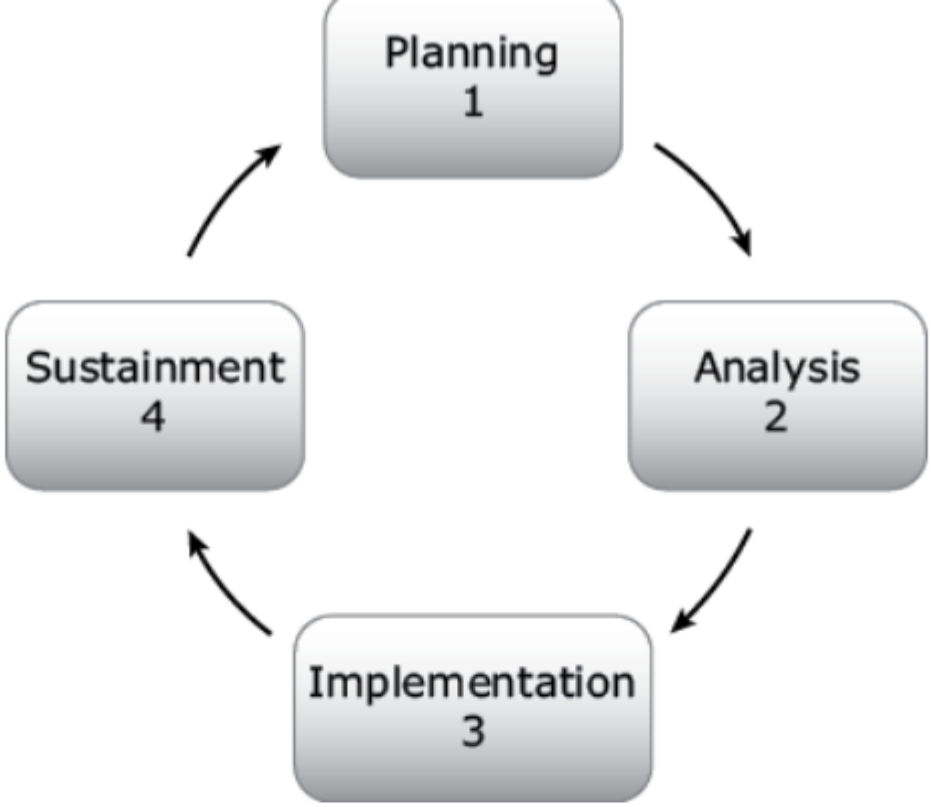


Figure 4.1. Systems development life cycle for health systems solutions

▶ Expand/collapse Extended Description

Planning Phase

This phase of the SDLC involves developing the appropriate IT-based approaches to support the organization's strategies. This stage is a precursor to further work that will lead to effective implementation of health system solutions. Stakeholders such as the chief information officer (CIO) develop IT-based plans for the implementation of solutions to enable the organization to be viable and competitive. The planning stage encompasses activities such as surveying the market for viable technology solutions and evaluating technology partners. High-level letters of inquiry for the vendor market called requests for information (RFIs) and detailed proposals for interested vendors called requests for proposals (RFPs) are initiated. During this stage it is also critical for organizations to identify sponsors to align the goals of technology implementation with the strategies of the organization and provide appropriate financial support. Technology leaders work with the organization to develop a technology plan to meet the business's needs; this determines whether solutions involve custom application development, the use of vendor partners, or a combination approach. Contracting, finalization of vendor partnerships, implementation plans, measurable outcomes, and project program cost estimates are finalized during this phase.

Analysis Phase

If sponsorship-level organizational support is eminent during the planning phase, leaders might encourage the analysis phase to begin in tandem. This phase of the life cycle includes collecting information about existing data requirements, functional and technical requirements, and business workflows. Data regarding types of end users who will use the technology is also collected during this time (Ozkaynak et al. 2021). This is also the time in the implementation process where it is important to determine what type of

hardware is needed. Questions for end users regarding the need for specific types of organizational computers, printers, handheld devices, kiosks, or bar code devices are discussed in this phase of the SDLC. The data gathered during this phase are turned into meaningful information to support building the system solutions, determining the user security protocol requirements, and instantiating best practice business workflows. Two other critical elements of effective implementation are addressed during this time: strategies to train users and strategies to ensure support models for both users and long-term technology support after the technology is implemented.

Implementation Phase

Implementation as a distinct phase of the SDLC involves building or developing, testing, and implementing the solution, whether it is software, hardware and devices, network upgrades, or new data sources. There are several steps involved in implementing technology solutions. Although presented in a linear fashion, some of the steps might overlap in the interest of expediting project and program timelines. The important emphasis is that each step is critical and that there are dependencies between certain steps. The amount of time and work allocated for each of these implementation steps is dependent on the type of solution being implemented, the goals of the implementation, and the resource availability of implementation teams and vendor resources.

Software solution configuration differs depending on whether the software being implemented is a software package, purchased from a vendor such as an EHR or a custom-developed solution such as a mobile health application developed by programmers internal to the health system. Configuration means doing the work to set parameters within the computer code of the system. This work occurs whether the software is a vendor solution or a custom-developed product. These configuration settings determine how the software functions for the end users. For example, when clinical order sets and decision support tools are being implemented in an EHR, configuration parameters must be entered into the software that determine which orders will be presented as groups to the clinicians, which types of orders will execute decision support roles in the software, and critical lab values that will determine how and what clinical decision support information is presented to a physician. Some configuration decisions happen within the implementation team; decisions that affect organizational policy are elevated to a governance group such as the medical executive committee for policy setting. Systems analysts, project managers, and team leads are responsible for discerning which configuration settings are addressed directly and which require executive sponsor input.

The **system building and development** step of implementation involves gathering and verifying data needed to input into software or hardware for the system to function. For example, in a vendor-created software package solution such as a human resource management system, a table of all existing employees must be built to enable the system to perform critical functions such as payroll and compensation. If a new cardiology camera system is implemented in 10 outpatient cardiology sites of a health system, the software that runs the cardiology cameras application must be loaded with site-specific data such as staff, procedures, and diagnostic information. In some cases, these data elements are input into the system by hand; in other cases, CSV or Microsoft Excel files with the data are uploaded into the new system. Either way, the implementation team must assure that the data are accurate before they are built into the new system.

If there is a business case to justify the cost, some sites choose to move records digitally or manually from old software applications into new software applications. This process is referred to as **data conversion**. Activities related to bringing data from an old platform into a new platform are costly. Prior to developing physical methods to move

data, implementation teams must develop a data quality strategy to ensure appropriate levels of quality in the data records that are being moved into a new platform. In some cases, individuals serve as abstractors to manually review subsets of data in the old system and input a subset of data into the new platform during the implementation phase. Sometimes data conversion involves moving digital data from one technology platform to a new platform; in other cases, data conversion involves inputting data from paper charts into a new platform. For example, the chart abstractors might review 2,000 digital health records of patients in an existing ambulatory EHR, extracting a subset of clinical and administrative data elements from each record for inclusion in a new EHR software implementation. The determination of these data elements and financial sponsorship for this process is approved by project or program sponsors. Sponsors approve this work based on the idea that many software systems will be easier to adopt if a subset of data is available when the system is put into productive use in the business.

Before solutions are introduced into business environments they must be thoroughly tested. **Testing** is paramount when implementing technology platforms of any type whether they are devices, new reporting capabilities for end users, or an entirely new transactional system such as a new EHR, supply chain management system, or human resource management system. Testing involves manipulating data that have been built into the system prior to turning the system on or going live for productive use in the business. Testing has three stages: unit testing, integration testing, and scalability or volume testing. In the unit testing and integration testing phases, data are manipulated using test plans developed in concert with end users. Test plans duplicate critical business processes such as admitting a client into a health facility or finalizing the payment of an insurance claim. Persons performing testing document any irregularities seen in the way the software performs or, in the case of testing, advise on the way that the device performs when replicating critical business functions. Unit testing involves ensuring that the device or software system performs as it should. Integration testing is the phase where data interfaces are duplicated, and a system is tested based on its capability to perform while taking data from another system. The last phase of testing, scalability or volume testing, is often done in an automated fashion. In smaller organizations without access to automated testing tools, multiple individuals log into platforms or use devices at the same time to simulate the action of multiple end users. Many vendors selling software to providers or insurance organizations have fully developed test plans and are a resource to support testing during the implementation phase.

When teams are implementing software solutions, end users require reporting and other data analytic functions such as visualization dashboards to be part of the solution. During the implementation phase, it is important for teams to test and verify the existing reporting functions and analytics functions of the solution being implemented. **Analytics and reporting verification** is a critical process and a formal step in the implementation plan. During this phase, there is also an opportunity for technology or business staff to be trained in reporting and data analytics platforms. Determining the reporting capabilities of the solution and verifying which reports will be used on a sustained basis should be completed before the solution is turned over for productive use to end users.

Health-related technology solutions rely on infrastructure that is sometimes unavailable. Examples of infrastructure include an internet connection, an internal business network, and computers with software and databases called servers that allow multiple end users to use a software solution such as an EHR. A critical step of the implementation phase of the SDLC is **downtime and disaster planning verification.** The strategy to address planned and or unplanned downtime and disasters is a topic for organizational governance and sponsors. Determining the level of financial resources appropriate to address managing the risks of downtime and disasters is a responsibility of

executive sponsors. This discussion is part of an organization's risk management plan for most health provider insurance or financial organizations who are handling health data. Project and program managers begin these discussions during the planning phase of the SDLC and ensure that the work involved to deliver feasible downtime and disaster solutions is completed during the implementation phase prior to the system going live for productive use. Supporting servers and databases that provide enterprise solutions for purposes of downtime is a high-level technology planning function.

Implementation teams work directly with management in operations ensuring that many types of operational support processes are in place to prepare for downtime. Some processes ensure that business operations can run when technology platforms are unavailable. Data in hospital systems must be managed in such a way that only those with approved access can view, append, or download the data. This concept refers to **data level security**, which is a set of technologies and processes that ensure appropriate technology platform and data access to approved users.. For example, does the organization have the capability to print out forms critical to the business from a personal computer that is not part of a network and that could potentially be compromised by unapproved users' attempts to gain access to a hospital network? Another example involves the implementation team working with supervisors in physicians' offices or clinics to set up processes where patient schedules are printed out at the end of every shift and secured in physically protected locations in the case of unexpected downtime. Perhaps a computer not on the enterprise network is in the supply area of the operating room and loaded with critical data files needed to set up basic surgical supplies in the case of the potential downtime event affecting the supply chain management and EHR systems. Due to the critical nature of some of the technology being implemented, program managers are encouraged to be creative in working with the operational colleagues to develop reasonable downtime solutions prior to the implementation of the system. Involving implementation resources such as vendors' staff or experienced implementation consultants is helpful in the design of downtime processes for critical systems such as EHRs (Gecomo et al. 2020). The implementation of any type of technology-based solution involves different levels of **training**. In this context, training means the knowledge transfer from experts to learners of how to use the technology as well as the new processes defined by the implementation of new or revised technology. Some training, often obtained through vendors, is needed for members of the implementation and support teams to learn the necessary technical skills to implement and support the solution. Strategies to ensure that the implementation team is adequately trained should be developed during the planning phase of the SDLC. The reality is that many implementation staff continue to obtain critical skills during the implementation of the new technology platform. The other aspect of training that occurs during the implementation phase is the training of end users. Depending on the type and number of end users, different strategies for training are used during the implementation phase. These include online training, pertinent newsletters about critical operational changes happening because of technology implementation, in-person sessions, synchronous online sessions, and informal modalities such as placing training devices and implementation training staff in physician lounges. Whatever the modality or strategy for training, it is critical to track competencies gained in not only utilizing the new technology, but new operational procedures instantiated because of a technology implementation.

Documentation involves ensuring that the data, technology decisions, support processes, and training methodologies used during implementation are digitally secured for future use. Every member of an implementation team has the responsibility to complete appropriate documentation related to the tasks that they execute during the implementation. Documentation is not only critical for supporting the system after going

live, it is also an imperative from a risk management perspective. Many health providers, insurance companies, and durable medical equipment companies are audited on a routine basis. The ability to produce accurate and timely documentation about systems used for daily business processes is a requirement not only for internal audits, but in many cases maintaining programmatic or organizational accreditation from external agencies. It is the responsibility of the project and program manager to ensure that documentation is completed during the implementation phase. Another aspect of project and program documentation involves communicating and clarifying how stakeholders plan to evaluate the performance of the technology being implemented and clinical and business performance planned as a result of the implementation.

Once implementation activities are completed, technology solutions are ready to be used in the business on a day-to-day basis. The technical, people, and process activities needed to begin the use of new systems and solutions are referred to as **conversion**. Once a new application has been turned on or new mobile devices are activated, the implementation team has responsibility for a fixed period to ensure that technologies and end users are stabilized using the new solutions in the business. This is referred to as **post-conversion support**. For example, if a new materials supply chain application is being implemented in a hospital, the implementation team will likely set up a manager first line help desk during the first two to three weeks of conversion where persons in the materials management and inpatient units and inpatient clinics can call directly for support involving any issues with the new application. Following up on these issues might involve implementation team members physically visiting a nursing unit where issues with pharmacy inventory are occurring, presumably as a result of the new system implementation. Working directly in tandem with end users, implementation team staff and any staff designated to support the system on an ongoing basis perform troubleshooting functions to ensure that critical deficiencies in technology, process clarity, or training are addressed. During this time, it is critical for project and program managers to collect data about post-conversion support and use these data in planning sustainable support processes.

Table 4.2 presents some examples of how some of the steps of the implementation phase might look in a real-world setting.

Summary of Substeps in the Implementation Phase

- Solution configuration
- System build and development
- Testing
- Analytics and reporting verification
- Downtime and disaster plan verification
- Training
- Documentation
- Conversion and post-conversion support

Sustainment Phase

Once an application, new reporting capabilities, new devices, or any type of new technology solution is converted and placed into productive day-to-day use, there is a commitment on the part of the organization to maintain support and enhance the technology solution. People, processes, technologies, and associated resources needed to ensure that solutions are managed effectively on an ongoing basis make up the **sustainment** phase of the SDLC. Sustainment is broken down into three areas: support, maintenance, and upgrades. Day-to-day support technologies post-implementation are

broken down into three levels of support: First-level support or level one means providing direct support to end users; second-level support involves analysts troubleshooting issues that have been determined to be technology related; and third-level support involves escalating issues that cannot be addressed on site to a vendor or consultant resource.

An important element of sustainment Is the superuser. A **superuser** is an individual who has business operations knowledge and whose part-time or full-time role is to receive additional training and support and maintain technologies during and after implementation. Planning the financial support reporting relationships and role for a superuser or a group of superusers is a topic discussed during the planning phase with sponsors and executive leadership. Some of the crucial activities that superusers address are related to nontechnical level one support issues, additional training of end users, writing and running reports and analytics dashboards, and initial unit testing and review of new products eligible for implementation upgrades. Superusers, some level of technology support, and end users are also often involved in maintaining systems post-implementation. For example, this might involve updating new Current Procedural Terminology (CPT) codes into the EHR on an annual basis. Upgrades and updates to technology are a requirement of most vendors. When the business determines that an update or upgrade is necessary and appropriate, the SDLC begins again with planning the upgrade and allocating the appropriate resources to move into the analysis and implementation phases.

Table 4.2. Analysis of two sample implementation phase milestones: EHR software implementation in an ambulatory clinic

Milestone	Tasks in project plan	Key stakeholder decisions
Complete configuration of orders and documentation decisions	Provider problem list management policy Determination of number and type of visit documentation templates Open and closed orders policies for providers	How do primary care and specialist clinicians manage the patient problem list? Which templates will be required for completing an ambulatory visit? Which templates are appropriate for specialist visits? How long will orders remain open before action is taken to complete an order in the clinic? What will the system parameter be set to in terms of number of hours?

Finalize downtime plan for telemedicine visits	Write auto-generated summary report printing out telemedicine visit requests Write office-specific downtime procedure for telemedicine visit interruption Document and train administrative and provider staff on telemedicine downtime procedure and patient callbacks Develop marketing materials and auto office phone messaging appropriate for patient communications regarding interrupted telemedicine services	What is the appropriate medical process for an interrupted electronic visit(s) and how will patient care be provided via telephone call requesting office visit?

Balancing Act: Cost, Quality, and Time

Planning, implementing, and sustaining technology-based transformational platforms involves a delicate balance of prioritizing and managing cost, quality, and time (figure 4.2). One of the best ways to ensure that implementations are effective in meeting their stated organizational goals while balancing these three factors is to assure that stakeholders are in alignment. In other words, the competing demands of executive stakeholders in any organization can be complex. During the planning phases of the SDLC, it is important to facilitate constructive and clear conversations about balancing cost and quality and agreeing on project and program goals that are supported by all executive stakeholders. During the planning phase, it is also important to be realistic about communicating the expectations of an implementation schedule to executive stakeholders. Vendors who implement systems and technologies all over the world or the region are an important resource, helping organizations to process realistic estimates of high-level timelines for implementation and communicate these timelines to executive leadership. Time spent working through executive stakeholder alignment and goal setting in the planning phase will pay off in the form of an efficient and effective implementation phase.

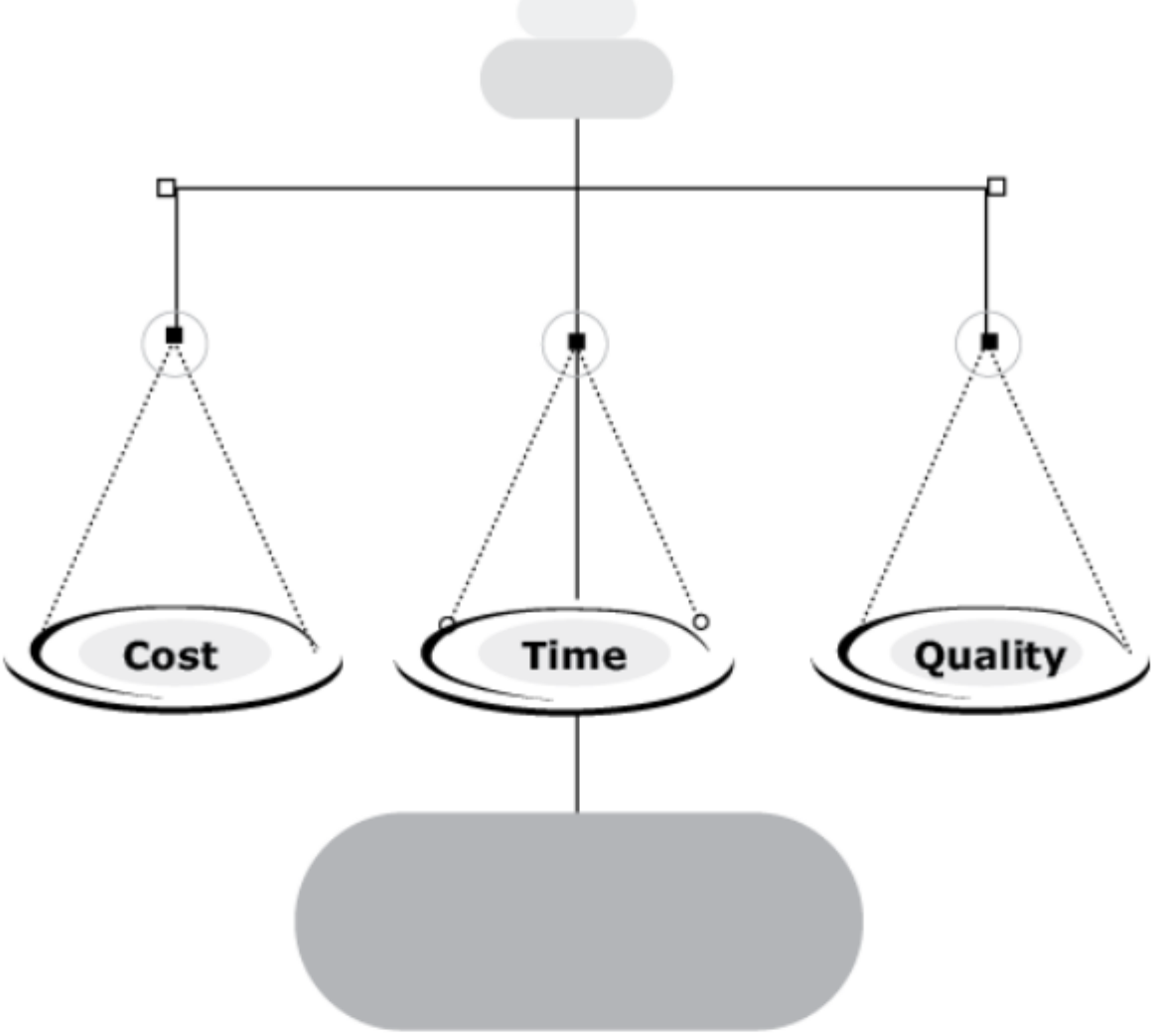


Figure 4.2. Balancing the triad during health systems implementation

▶ Expand/collapse Extended Description

Case Study: Human Resource Management System (HRIS) Upgrade with RFID Badges

The HRIS of a community hospital is in the planning stages of a major software upgrade. Given pandemic-related financial effects on the health system and increased competitive pressures in the region, the IT budget for this upgrade has been reduced by 10 percent. The original budget included the implementation of radio frequency (RF) capabilities for staff associate badges for the hospital, including increased security measures by way of physical security, pictures on badges and RFID swiping to track staff operations. Given budget cuts, the chief human resources officer (CHRO) plans to update the executive committee suggesting that this year's upgrade of the HRIS system be implemented without RFID badge functionality. The upgrade must move from planning to implementation in the next three

months to meet deadlines for changes in compensation compliance that require the new software to be in place at the start of the new fiscal year.

Technology Considerations and Implementation

The health sector of the economy involves organizations of all types. Health provider organizations include both for-profit and nonprofit systems of all sizes spanning the gamut of sizable academic medicine institutions through including community health systems and independent physician provider practices in urban and rural communities. Provider organizations also include home care agencies, long-term care and assisted living facilities, as well as addiction rehabilitation facilities and general rehabilitation organizations. Insurance entities include government-based insurance programs such as Medicaid in addition to private insurance companies. The health sector also includes several organizations that function daily to provide services to consumers such as medical equipment companies, staffing agencies, and pharmaceutical organizations. Regardless of the nature or mission of the organization, a commonality in all organizations is the need to use technology-based solutions effectively to deliver quality products and services, create data to inform better decision-making, and excel in customer service.

Types of Technology Solutions

Technology platforms systems and components are varied and complex. In addition to a variety of health organizations, there are also a variety of types of technologies being planned, implemented, and sustained in any of these organizations at any point in time. The best practices approaches discussed today involving people and process are applied to a host of different technologies. These technologies fall into common categories for discussion purposes but are most often implemented in combinations of technology solutions based on the business needs and goals of the organization.

Software Implementation

Whether the computer code is custom developed by a health organization, a consultant, or a university or purchased from a vendor, a great deal of activity in the health economy regarding implementation involves implementing software solutions. Software solutions align with the business processes of organizations. For example, EHR solutions align with health providers' needs to deliver healthcare in settings such as inpatient settings, ambulatory settings, virtual medicine settings like telemedicine, or in someone's home. Best practices in stakeholder management and utilization of the SDLC ensure that the clinical financial and customer service goals of an EHR are met or exceeded as vendor negotiations are finalized and a program is kicked off to automate hospitals, clinics, and provider offices in a community health system. Likewise, transactional software solutions automate processes in accounting, supply chain management, planning, marketing, and human resources talent management in virtually every health organization. In other words, a dentist office implementing a new accounting and insurance system will utilize the same SDLC approach to implement a new vendor-based solution as a durable medical equipment company updating a new system to manage customers and provide invoicing.

Updates and Upgrades

One of the universal complexities involved in any software or hardware platform is the requirement for constant updates and upgrades. Updates typically happen in a shorter and more expeditious fashion. For example, a system administrator managing 5,000 clinical

devices at a health system might automatically design, test, and push out a security update in a matter of four hours. The SLDC in this case is short but involves planning, testing, deploying, and ensuring that the security patch is complete. On the other hand, major software packages such as EHRs or HRISs require major upgrades, typically within 24- to 36-month cycles; many of which are determined by contract requirements with vendors. Upgrades of major software packages such as EHRs, supply chain management systems, and enterprise resource management systems are typically formidable projects in size and scope and must be executed through all the phases of the SDLC beginning with planning. It is not uncommon for the upgrade of a software package such as an EHR system that automates 400 physician practices in a health system to include a total scope of work exceeding 30,000 to 40,000 people hours.

Hardware

Different types of computer hardware are used in projects and programs that are implemented in health organizations every day. Examples include several new database servers being upgraded in a health system to support the EHR, providing services to 40,000 users. Other examples of hardware implemented in health organizations include desktop devices, printers, RFID scanners, mobile devices, kiosks for customer check-in and checkout, and biometric monitors used to support both inpatient and ambulatory physiologic monitoring of patients. In some cases, hardware implementation happens within the context of a software project or program. In other cases, hardware implementation or hardware upgrades constitute their own projects or programs. Regardless of how hardware solutions are implemented, it is critical that the hardware itself and the operating systems controlling that hardware follow an SDLC process like that used with software implementation. For example, before deployment of mobile devices to providers in a health system setting, the devices themselves must be thoroughly planned in terms of asset management, cost, and ownership, appropriately configured for software, and tested appropriately before being turned over to end users. There must also be a sustainment program put in place to support end users who are using the devices.

Services

Implementation in health settings also include services. Services involve the use of internal or external technologies, individuals, and expertise to provide value to an organization. For example, a small community health system with limited IT staff resources might choose to use the services of a consultant provider such as a managed service provider to implement secure networks, and implement and support network email and other day-to-day software and hardware solutions in their hospital clinics and provider offices. The implementation of this service model to support the community health system with secure and reliable technology services follows an SDLC implementation approach like implementing any other type of solution. Similarly, the CEO of a large medical facility may decide that the organization can use assistance from best-in-class vendors regarding the use of their services to process data into information and predictive models related to clinical care and administrative excellence. The CEO could initiate a planning project designed to explore vendors in the market who offer analytics services and best-in-class solutions to address this need. The SDLC is started, and if moved forward, the service models and any associated data models and dashboards will be planned, analyzed, implemented, and sustained within the organization.

Value-Based Technology Implementation

Health organizations that achieve a level of strategic success systematically implement appropriate technological solutions not only for day-to-day business leads, but also to

address their competitive position (Upadhyay and Opoku-Agyeman 2021). Significant financial investments in areas such as EHRs, human resource and supply chain systems, and marketing platforms must be implemented in a way that derives continued value for an organization (Chopra and Bonello 2019). To a hospital system, physician practice, insurance company, medical equipment company, or long-term care facility, the value of the technology investment is measured over time by comparing the cost of acquiring, implementing, and managing the technology. Different stakeholders perceive the value of technology investments and the new processes they define in different ways. For example, the CFO might evaluate whether the implementation of a system reduces administrative expenses, whereas physicians evaluate value based on the ability for chronically ill patients to effectively manage their medications (Mayzell et al. 2021). Regardless of perceptions of value by any stakeholder, generally the implementation of technology considers costs in the context of the ability of the technology to adapt effectively to changes internal and external to an organization.

Implementation Costs

The allocation and support for financial resources required to implement health system solutions of any type come from executive organizational governance organizations and are administered through project and program sponsors. Implementation project and program managers are responsible for estimating, understanding, and managing the costs of implementation. Organizations manage and obtain financial resources in different ways to support different types and levels of implementation activity. For this reason, it is critical for implementation project and program leaders to understand financial standards within the organization such as definitions of capital and operational costs. **Capital costs** associated with technology implementation are costs that can be depreciated by the organization. Capital costs are often defined in health organizations by the size of the dollar investment. Capital costs are also typically planned and managed in a multiyear planning cycle. **Operational costs** represent financial resources paid at the time that the dollars are spent. For example, purchasing five servers to support the implementation of an HRIS is likely a capital cost. The salaries of employees in part used to implement those servers might be a capital cost or totally managed as an operational expense. Conversely, salaries that are used for training during the implementation must always be categorized as operating expenses. These parameters are managed in detail by financial professionals within the organization reporting through the function of the CFO. In the planning stage of the SDLC, it is important for implementation leaders to work through cost estimates for the implementation and detailed annual estimates of financial responsibility for sustaining the technology long-term.

Implementing Adaptable Solutions

It is evident that both health organizations and technology solutions are varied. Adding to this complexity, twenty-first-century competitive pressures in an information economy make it more critical than ever for health organizations of any type to be highly adaptable. Technology solutions are more component based and, in many cases, more cost-effective than in decades past. Implementation teams have the responsibility to implement technology-based solutions for health organizations in ways that allow solutions to not simply survive but to adapt, thrive, innovate, and compete. Implementing **adaptable solutions** means ensuring that systems are implemented in the least proprietary, most cost-effective, and most user-empowered way possible to facilitate organizational benefits (Colicchio et al. 2019). Componentized web solutions, powerful end user analytics dashboards, relational databases with accessible data, and mobile computing add-ons are some of the technology capabilities that enable adaptable solutions.

For example, registries and dashboards within digital health platforms implementing adaptable methods enabled system administrators in health systems to configure tools for the timely and appropriate management of the coronavirus in health systems (Jose et al. 2021). In another example, order sets in an EHR are adopted to address alternative solutions for opioid prescribing in an emergency department setting (Netley et al. 2020). In each instance, standard software tools within a package system that has been implemented effectively are used to address evolving needs of the health organization. Technology solutions implemented and converted into productive use the first time provide a rich asset to the organization that it can continue to use creatively on an ongoing basis. Realizing these platforms in the most efficient way involves a series of principles based on adaptability.

Summary Principles Implementing Adaptable Solutions

- Explore the most cost-effective and flexible way to implement software including contracting hosted software-as-a-service licensing as opposed to exclusively using traditional license models.
- Encourage vendors and custom development teams to build end-user-friendly system utilities enabling changes in basic workflow-driven software configuration points and data maintenance.
- Train superusers in reporting and data analytics functions.
- Avoid the use of proprietary software, hardware, and data analytics solutions as much as possible.
- Implement rigorous quality assurance programs reviewing the quality of post-implementation documentation.
- Implement rigorous system governance models that include clear guidelines for end users and superuser roles.
- Participate actively in vendor engagement activities such as user groups and development committees.
- Develop executive governance-level committee connections between risk management, financial management, and technology strategy planning committees.

Case Study: Implementing a Tethered Personal Health Record

An academic medical center provider is in the planning stages of expanding the PHR module tethered to their EHR platform. The tethered portal was implemented across all physician practices three years ago to meet the basic messaging requirements for Meaningful Use phase 3. In alignment with the strategic plan for the medical center, the executive technology committee authorizes the use of resources to improve population health management, deliver additional online visit capabilities, and reduce administrative costs associated with scheduling and prescription refill management across the health system. Additional functionality and rigorous campaigns encouraging the adoption of the tethered PHR is the basis for this strategy.

While investigating the operating costs annually of hosting one patient on the portal, the implementation program manager determines this cost to be $1.50 per patient per year. Given the fact that the whole system has more

than 3 million people in their database, subscription cost estimates to implement the portal are substantial. A team of implementation leaders has been working on a plan with financial analysts from the CFO's office to make a recommendation to the executive committee regarding online bill payment as the first project of the PHR implementation program. The reason for this is that the health system finance committee estimates that if 20 percent of patients pay bills online, the net annual savings in processing paper statements and increasing cash flow in accounts receivable pays for the subscription costs of the overall PHR.

Chapter 4 Review Exercises

Discussion

1. Describe two examples of operational costs that might be incurred during the upgrade of an EHR system in a hospital.
2. Outline three functions that superusers serve in the implementation process.
3. How can the project manager implementing a PHR system ensure that the clinical functionality is as adaptable as possible?
4. Determine whether the purchase of three computer servers to run a database and support 2,000 users for the implementation of a new supply chain system is a capital or operational expense.
5. List an example of a service that might be purchased by the CEO of a small hospital when implementing a new HRIS.
6. Explain the role of a super user in the implementation of a software solution.
7. Referring to Case 4.1 (HRIS) system upgrade with RFID Badges. Describe the balancing act involving cost, time and quality in this implementation scenario.

Answer the following questions.

8. Which statement best describes whether a cost related to implementing a new supply ordering system in the pharmacy is capital or operational in nature?
   a. whether the cost is related to software or hardware
   b. whether the cost is planned at least one year before the start of the implementation
   c. whether the cost involves dollars that the organization cannot depreciate
   d. None of the above
9. Which definition best describes the concept of a system conversion?
   a. Creating appropriate end user access to data in the new platform
   b. Defects found in the software during testing
   c. Technical, people and processes needed to begin using a new technology platform
   d. All of the above
10. Which phrase is not a core principle in implementing technology solutions that are adaptable?
    a. Implement rigorous system governance models that include clear guidelines for end-users and super user roles
    b. Use proprietary software, hardware, and data analytics solutions as much as possible

c. Develop executive governance level committee connections between risk management, financial management and technology strategy planning committees
d. Train super users in reporting and data analytics functions

REFERENCES

Abdekhoda, M., A. Dehnad, and J. Zarei. 2019. Determinant factors in applying electronic medical records in healthcare. ***Eastern Mediterranean Health Journal*** 25(1):24–33. https://doi.org/10.26719/emhj.18.007.

AHIMA. 2002. "Who we are." https://ahima.org/who-we-are/about-us/.

AMIA. 2022. "Why informatics." https://amia.org/about-amia/why-informatics.

Andrews, S. M., K. A. Porter, D. B. Bailey, Jr., and H. L. Peay. 2022. Preparing newborn screening for the future: A collaborative stakeholder engagement exploring challenges and opportunities to modernizing the newborn screening system. ***BMC Pediatrics*** 22(1):1–13. https://doi.org/10.1186/s12887-021-03035-x.

Bangash, H., L. Pencille, J. H. Gundelach, A. Makkawy, J. Sutton, L. Makkawy, O. Dikilitas, S. Kopecky, R. Freimuth, P. J. Caraballo, and I. J. Kullo. 2020. An implementation science framework to develop a clinical decision support tool for familial hypercholesterolemia. ***Journal of Personalized Medicine*** 10(3). https://doi.org/10.3390/jpm10030067.

Chopra, S. J., and J. Bonello. 2019. How to achieve a return on an EHR. ***Healthcare Financial Management***, Vol.73 (3), p.30.

Colicchio, T. K., D. Borbolla, V. D. Colicchio, D. L. Scammon, G. Del Fiol, J. C. Facelli, W. A. Bowes, III, and S. P. Narus. 2019. Looking behind the curtain: Identifying factors contributing to changes on care outcomes during a large commercial EHR implementation. ***EGEMS (Generating Evidence & Methods to Improve Patient Outcomes)*** 7(1):1–12. https://doi.org/10.5334/egems.269.

Gecomo, J. G., A. Klopp, and M. Rouse. 2020. Implementation of an evidence-based electronic health record (EHR) downtime readiness and recovery plan. ***Online Journal of Nursing Informatics*** 24(1):5.

Georgiou, A., F. Magrabi, H. Hyppönen, Z. S.-Y. Wong, P. Nykänen, P. J. Scott, E. Ammenwerth, and M. Rigby. 2018. The safe and effective use of shared data underpinned by stakeholder engagement and evaluation practice: A position paper from the IMIA Technology Assessment & Quality Development in Health Informatics Working Group and EFMI Working Group for Assessment of Health Information Systems. Yearbook of Medical Informatics 27(1):25–28. https://doi.org/10.1055/s-0038-1641194.

Hagland, M. 2017. Value-based healthcare and IT: HIT leaders work out new strategies. ***Healthcare Informatics***:79(6)

Halal, W. E., J. Garretson, and O. Davies. 2022. Updating strategy for a high-tech world: Constant change from the bottom up and the outside in. ***Foresight*** 24(1):37–54. https://doi.org/10.1108/FS-02-2021-0051.

Hartley, D. M., S. Jonas, D. Grossoehme, A. Kelly, C. Dodds, S. M. Alford, E. Shenkman, J. Simmons, L. C. Bailey, H. Razzaghi, L. H. Utidjian, J. McCafferty-Fernandez, F. S. Cole, J. Smallwood, L. N. Werk, and K. E. Walsh. 2020. Use of EHR-based pediatric quality measures: Views of health system leaders and parents. ***American Journal of Medical Quality*** 35(2):177–185. https://doi.org/10.1177/1062860619850322.

Janssen, A., C. Donnelly, E. Elder, N. Pathmanathan, and T. Shaw. 2021. Electronic medical record implementation in tertiary care: Factors influencing adoption of an electronic medical record in a cancer centre. ***BMC Health Services Research*** 21(1):1–9. https://doi.org/10.1186/s12913-020-06015-6.

Kalayou, M. H., B. F. Endehabtu, and B. Tilahun. 2020. The applicability of the modified Technology Acceptance Model (TAM) on the sustainable adoption of eHealth systems in resource-limited settings. ***Journal of Multidisciplinary Healthcare*** 13:1827. https://doi.org/10.2147/JMDH.S284973.

Kettlewell, J., J. Phillips, K. Radford, and R. dasNair. 2018. Informing evaluation of a smartphone application for people with acquired brain injury: A stakeholder engagement study. ***BMC Medical Informatics and Decision Making*** 18(1):1–13. https://doi.org/10.1186/s12911-018-0611-0.

Lai, P. C. 2017. The literature review of technology adoption models and theories for the novelty technology. ***Journal of Information Systems & Technology Management*** 14(1):21. https://doi.org/10.4301/S1807-17752017000100002.

Land, S. K., D. B. Smith, and J. W. Walz. 2012. ***Practical Support for Lean Six Sigma Software Process Definition: Using IEEE Software Engineering Standards***. New York: John Wiley & Sons.

Mayzell, G., S. Berger, and D. McKinley. 2021. Physicians and finance: Managing the differences is critical to building clinical value: For finance and clinical leaders to work together effectively, they must not only understand but also acknowledge the benefits to be gained from their different perspectives on what constitutes success. ***Healthcare Financial Management*** 75(5) Pg 34.

Radack, S. 2022. ***The System Development Life Cycle (SDLC).*** Gaithersburg, MD: National Institute of Standards and Technology.

Rumball-Smith, J., K. Ross, and D. W. Bates. 2020. Late adopters of the electronic health record should move now. ***BMJ Quality & Safety*** 29(3):238–40. https://doi.org/10.1136/bmjqs-2019-010002.

Soh, F., and V. Grover. 2020. Effect of release timing of app innovations based on mobile platform innovations. ***Journal of Management Information Systems*** 37(4):957–987. https://doi.org/10.1080/07421222.2020.1831763.

Tsai, C. H., A. Eghdam, N. Davoody, G. Wright, S. Flowerday, and S. Koch. 2020. Effects of electronic health record implementation and barriers to adoption and use: A scoping review and qualitative analysis of the content. ***Life (2075-1729)*** 10(12):327. https://doi.org/10.3390/life10120327.

Unhelkar, B. 2016. ***The Art of Agile Practice: A Composite Approach for Projects and Organizations***. Boca Raton, FL: CRC Press.

Upadhyay, S., and W. Opoku-Agyeman. 2021. Factors that determine comprehensive categorical classification of EHR implementation levels. ***Health Services Insights (14)***:1–10. https://doi.org/10.1177/11786329211024788.

Chapter

5

Understanding Databases

By Jacqueline Woeppel, ScD, MBA

Learning Objectives

- Identify the difference in flat data files versus relational databases
- Apply basic SQL commands to select data for reporting and analysis
- Utilize a data dictionary to understand data attributes
- Discover the role of data modeling in database maintenance and design
- Understand the various types of data resources
- Understand importance of Data Marts and Data Warehouse

KEY TERMS

Cardinality
Clinical data warehouse
Column-delimited data
Comma-separated values (CSV)
Data dictionary
Data flow diagram
Data mart
Data model
Data table
Data warehouse
Database
Database management system (DBMS)
Decision support databases
Diagram 0
Entity relationship diagram (ERD)
Extract, transform, and load (ETL)

Field
Flat file
Foreign key
Join
Method
Normalization
Not Only Standard Query Language (NoSQL)
Object-oriented databases
Ontology
PostgreSQL
Primary key
Record
Relational databases
Select query
Structured Query Language (SQL)

Healthcare is a data-rich industry. On the operational and financial side of healthcare, claims are generated and submitted for payment. Payment transactions are then sent back to the provider. On the clinical side of healthcare, countless diagnostic tests are performed and the results of those tests are provided back to practitioners and providers. Both operational and clinical data are used for healthcare services research and process improvement projects. The introduction of the electronic health record (EHR) and sophisticated electronic medical record systems have increased the amount of data available in the healthcare setting exponentially. Governmental programs such as The Centers for Medicare and Medicaid Services (CMS) Quality Payment Program (QPP), Merit-based Incentive Payment Programs (MIPS), value-based purchasing programs and the formation of Accountable Care Organizations (ACOs) are driving healthcare providers to find ways to help improve the efficiency and effectiveness of the healthcare delivery system.

Healthcare systems continue to expand data utilization to help support population and public health (public health emergency (PHE) and pandemic) requirements. To meet these important goals, data must be organized to ensure any resulting analysis and reports are timely and accurate. Databases give foundational structure to the raw data and facilitate the manipulation of data to meet the evolving needs to achieve these goals.Database Terminology

Prior to studying the various types of databases and database management systems (DBMSs), it is important to have a common set of terminology to describe the various components of a database. Databases can be envisioned as a collection of data tables. The most common form of a data table found in practice is a worksheet within a spreadsheet.

Data tables include records or rows and fields or columns. For example, table 5.1 represents a table of patients and their associated appointments. The fields in this table represent data elements describing attributes of the appointment (provider, date, and time). A collection of fields that are related are placed in the same record. The records in this table represent the data describing the patient's appointment. Records in a data table may represent patient demographic information, attributes of a service provided to a patient, or contain diagnostic or procedural codereferences, such as the ICD-10-CM, ICD-10-PCS, and current procedural terminology (CPT).

Table 5.1. Patients and their appointments

PatientID	ProviderID	Appointmenth scope="col"ate	AppointmentTime
ABD239	SMI123	1/23/2023	9:00 a.m.
DIR235	SMI123	1/23/2023	9:30 a.m.
JKF764	SMI123	1/23/2023	10:00 a.m.

Sets of records include a common set of fields and those related to the same business purpose or focus are collected together into a data table. Common data tables found in healthcare include information on the following:

- Patient demographics
- Physician specialty and licensure
- Charge description master
- Patient services

A collection of data tables is called a **database**. Databases can have a variety of structures and may be created and maintained using a **database management system (DBMS)**. A DBMS provides a method for adding, updating, or deleting data and also supports methods for extracting data for various purposes to support organizational decision-making (for example, productivity ,quality measures and staffing requirements).

Types of Databases

To decide on the type of database that is most appropriate for storing data, the user must answer some key questions:

- How many records will the database contain (currently and in the foreseeable future)?
- How do the fields (variables) relate to each other?
- What database tools are available for use?
- How will the data be pulled or queried from the database?

The answers to these questions will drive the database type. There is typically not a right or wrong answer in database design. If the database is robust enough to hold all of the information required and extracting information is straightforward, then it should be considered a good quality design. The two most common types of data storage used in practice are flat files or data tables and databases. Databases can be further segmented into those that are relational or those that are not relational. Relational databases include tables that are related by common fields called keys. Nonrelational databases, including object-oriented databases, hierarchical databases, and **Not Only Standard Query Language (NoSQL)** databases, are used in healthcare because they allow flexibility in the storage and retrieval of non-structured data such as images and free-text note fields commonly found in EHRs. NoSQL use is growing today, driven by cloud computing and big data. Companies such as Oracle, Amazon, and Google are leading the way as alternatives to SQL-only database manipulation.

Flat Files

If the data structure is relatively simple, then a **flat file** or spreadsheet might be the right type of data storage. A flat file is a text file, usually delimited by a comma or tab, with one record found on each row. It has only one table of data. For instance, if a practice

manager wished to track the number of records coded by each employee, then the flat file displayed in table 5.2 is sufficient.

This structure allows the manager to see who is the most or least productive and can be expanded by adding additional columns to track more days. A series of flat files or individual spreadsheets could be used to track other information about the employees such as credentials or continuing education units (CEUs). If the number of additional attributes is small, then they may be added to the records displayed in table 5.2. If there is a large number of additional attributes, they can be stored in a second flat file.

Flat file databases may be stored in a number of formats. Traditional spreadsheet worksheets, such as those found in Microsoft Excel, are the most common format found in business applications. Flat files may also be stored in text files as columns of data or with the fields delimited by a character. The data in table 5.2 are **column-delimited data** (a flat file database format where information is stored in text files as columns of data). The same data are shown in table 5.3 as a **comma-separated value (CSV)** file (a flat file database format where fields are delimited by a comma). Column-delimited data must be accompanied by documentation, or a **data dictionary**, that lists the order and position of the variables so that the data may be interpreted properly. CSV files might include the variable names in the first row. If not, then they also require documentation to identify the variables in each column.

Flat data tables have an intuitive two-dimensional format of rows and columns. The number of rows and columns is limited only by the software tool used to store and analyze the data. If a spreadsheet program is used to store the data, then formulas may be embedded into the table structure and are stored as part of the data table, as one can do in Microsoft Excel. A flat data table can be sorted or filtered to select particular values of the variables.

Table 5.2. Flat file

Employee	Monday	Tuesday	Wednesday	Thursday	Friday	Weekly Total
Anne	4	9	6	6	5	30
Nicholas	6	10	7	3	6	32
Zach	4	7	9	8	4	32

Table 5.3. Comma-separated values

Employee, Monday, Tuesday, Wednesday, Thursday, Friday, Weekly Total
Anne,4,9,6,6,5,30
Nicholas,6,10,7,3,6,32
Zach,4,7,9,8,4,32

The first important limitation of storing data in flat tables is that the design does not define and maintain relationships between variables in separate tables. This limits the ability to cross-check values for data integrity with those variables stored in the same table. For instance, if the charge description master (CDM) includes CPT codes, the validation of the CPT code could only be accomplished by relating the CDM in one table to the CPT code table using the CPT code. The second limitation is that the same data element might need to be stored in multiple tables to allow for useful reporting. Data redundancy is not a good practice in database design and can cause serious data integrity issues when a variable is updated in one table, but not the others in which it is stored. If the tables used to store the various data elements need to be combined or must be

validated against each other to ensure data integrity, then a relational database is a better choice.

Relational Databases

Relational databases were first conceptualized by Edgar F. Codd. He proposed that data should be stored in a way that allowed users to query and analyze data without having to restructure it (Codd 1970, 377). As an example, a data analyst wished to create a report of the number of patients served by zip code and clinical department. The data reside in two tables. One holds the patient demographic information. The second holds the information about the patient's visit, including the clinical department. If the database is stored in flat files, then the clinical department must be added to the first file or the zip code must be added to the second file. If the database is relational, then the two tables can be joined or linked using the patient identifier as a common data element.

In a relational database, data with a common purpose, concept, or source are arranged into tables. The relationship between the tables is displayed in an **entity relationship diagram (ERD)**. An ERD is defined as "a specific type of data modeling used in conceptual data modeling and the logical-level modeling of relational databases" (AHIMA 2017). In essence, ERD visually illustrates the relationship between different entities such as patients and items. ERD may be used in industries such as engineering, business, and healthcare. Figure 5.1 displays a simple ERD relating a patient information table to a table containing the dates of service for each visit for all patients. This ERD shows that the patient information and visit tables are related by the PatientID variable.

Relational databases are structured in a way that helps ensure data integrity. Notice that in figure 5.1, one variable in each table has an icon of a key next to it. These are the primary keys in each of the tables. The **primary key** uniquely identifies the row in the database. PatientID uniquely identifies a row in the patient information table and AccountNumber uniquely identifies the row in the visits table. Properly defining the primary key field in a table prevents the addition of duplicate values of that field in the tables. A **foreign key** is a variable in one table that is a primary key in another table. In this example, PatientID is a foreign key in the visits table.

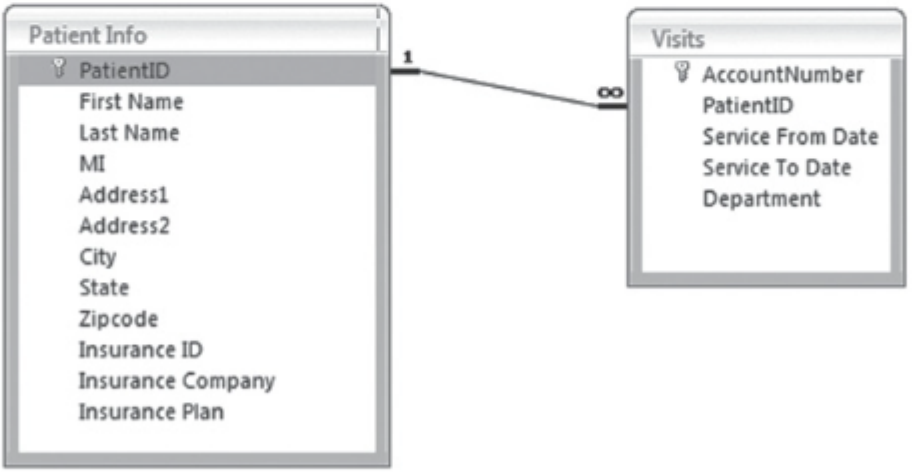


Figure 5.1. Entity relationship diagram

▶ Expand/collapse Extended Description

Used with permission from Microsoft.

The line extending in figure 5.1 from the PatientID variable in the patient information table to the PatientID variable in the visits table indicates that these two tables are related through the PatientID field. **Cardinality** represents the relationship between the two tables. Cardinality for each table refers to the number of elements in each table. Table 5.4 presents three types of relationship cardinality in relational databases:

1. One-to-one: Each row in one table relates to one and only one row in the other.
2. One-to-many: Each row in one table may relate to many rows in a second table; each row in the second table relates to only one row in the first table.
3. Many-to-many: Each row in one table may relate to many rows in a second table; each row in the second table may relate to many rows in the first table.

Carefully defining the variables in each table and the primary key fields for each table is the first step in ensuring data integrity. Data redundancy or duplication should be avoided in a relational database. The practice of **normalization** of a database prevents duplication of data elements and ensures the data all conform to a standard. There are three forms of normalization:

1. First Normal Form
   a. Eliminate repeating groups in individual tables
   b. Create a separate table for each set of related data
   c. Identify each set of related data with a primary key
2. Second Normal Form
   a. Create separate tables for sets of values that apply to multiple records
   b. Relate these tables with a foreign key
3. Third Normal Form
   a. Eliminate fields that do not depend on the key (Microsoft 2022).

Table 5.4. Cardinality

Cardinality	Table A	Table B	Explanation
One-to-one (1:1)	Patient ID Discharge status code	Discharge status code Definition	Each patient has only one discharge status code; each discharge status code has only one definition.
One-to-many (1:N or 1:∞)	Account number Patient name	Account number Diagnosis code sequence ICD-10-CM diagnosis code	Each account may have many associated diagnosis codes.
Many-to-many (N:N or ∞:∞)	Account number Bill type Bill date	Account number Bill type Bill date Payment amount	Each account may have multiple bill types and bill dates. Each bill type and bill date may have multiple payments.

Case Study: Normalizing Data Tables

In this case study, the flat data table included in table 5.5 will be converted to a relational database with a normalized form. When data are normalized, the data in one flat file may result in a number of tables.

The first normal form requires that the repeating groups be eliminated, separate tables be created for each set of related data, and each resulting table have a primary key. In table 5.5, the variables may be segmented into patient information (table 5.6) and service information (table 5.7).

The services table may be further broken down into a revenue code definition table, services by account number, and a CDM table. These are displayed in tables 5.7a, 5.7b, and 5.7c.

The final step in first-order normalization is to identify the primary key fields in each table. The primary key for the visits table is the account number. The primary key in table 5.7a is a combination of the account number and the CDM item. The primary key in table 5.7b is the CDM item, and, finally, the primary key in table 5.7c is the revenue code.

By eliminating repeating rows, the tables now actually conform to the second normalized form. Recall that the second normalized form requires that separate tables be created for sets of values that apply to multiple records and that the tables be related with a foreign key.

The ERD displayed in figure 5.2 shows the relationships between the normalized tables. Notice that the field names were converted to single words by inserting an underscore between the words in the original variable names. It is best practice to use single-word variable or field names and table names to simplify the syntax for writing queries and reports.

The final step in normalization is to check the third normalized form. To conform to the third normalized form, all fields that do not depend on a key (primary or foreign) must be eliminated. The tables in figure 5.2 do not include any fields that do not depend on the key fields and therefore the tables are completely normalized.

Object-Oriented Databases

Object-oriented databases are designed to handle data types beyond text and numbers. Flat file and relational databases were developed to store data that fits into rows and columns. Object-oriented databases may be used to store images or videos. In healthcare, object-oriented databases may be used to store images from an MRI or an x-ray, or even an audio file capturing the heartbeat heard during a prenatal ultrasound. The building blocks of an object-oriented database are the objects, as opposed to the tables in a relational database system.

An object-oriented database stores two types of information about the object. The first element is the data itself (audio clip, image, video file, and such). The second element stored describes how to use the data and is called the **method**. The expansion of EHRs requires the storage of many objects beyond the simple rows and columns of numbers and text that may be stored in relational databases.

Database Management Software

To access and manipulate a database, it must reside in a software tool. The choice of database software depends on the type of database, the size of the database, and the

complexity of the relationships between the database elements. Relational databases are found most often in practice. The most common relational database software applications include Microsoft Access, Oracle, MySQL, and Microsoft SQL Server. **PostgreSQL** is an object-relational database management system. In other words, PostgreSQL can manage relational databases with tables. It can also manage various types of objects, such as audio, video, and images.

Microsoft Access is typically run from a workstation and is appropriate for smaller department-specific applications. It is also useful for prototyping larger enterprise-wide database systems. Once the number of users expands and Access can no longer efficiently accommodate the database, the database might require a more sophisticated software application. Microsoft offers a wizard that allows users to convert an Access database to a SQL Server database (Microsoft 2021).

Table 5.5. Normalized data

MRN	Account number	Date of service	CDM item	Revenue code	Revenue code definition	HCPCS code	Units	Charge
123ABC	1	011011	L01	0300	LABORATORY OR LAB	36415	1	$55.75
123ABC	1	011011	L02	0300	LABORATORY OR LAB	84066	1	$51.50
123ABC	1	011011	L03	0301	LAB/CHEMISTRY	84153	1	$92.25
123ABC	1	011011	C01	0331	CHEMOTHER/INJ	96402	1	$168.50
123ABC	1	011011	D01	0636	DRUGS/DETAIL CODE	J9217	1	$507.77
987ZYX	2	010411	D02	0250	PHARMACY		6	$306.04
987ZYX	2	010411	S01	0270	MED-SUR SUPPLIES		1	$11.25
987ZYX	2	010411	S02	0270	MED-SUR SUPPLIES	A6257	1	$0.42
987ZYX	2	010411	L04	0305	LAB/HEMATOLOGY	85025	1	$107.00
987ZYX	2	010411	C02	0335	CHEMOTHERP-IV	96413	1	$495.00
987ZYX	2	010411	E13	0510	CLINIC	99213	1	$136.50
987ZYX	2	010411	D03	0636	DRUGS/DETAIL CODE	J2405	8	$50.65
987ZYX	2	010411	D04	0636	DRUGS/DETAIL CODE	J9305	100	$10,714.56
987ZYX	3	020611	S03	0270	MED-SUR SUPPLIES		1	$11.25
987ZYX	3	020611	S02	0270	MED-SUR SUPPLIES	A6257	1	$0.42
987ZYX	3	020611	E13	0510	CLINIC	99213	1	$136.50

Table 5.6. Patient information

Patient_ID	Account_Number	Service_Date
123ABC	1	01/10/2011
987ZYX	2	01/04/2011

987ZYX	3	02/06/2011

Table 5.7. Service information

Services						
Account number	**CDM item**	**Revenue code**	**Revenue code definition**	**HCPCS code**	**Units**	**Charge**
1	L01	0300	LABORATORY OR LAB	36415	1	$55.75
1	L02	0300	LABORATORY OR LAB	84066	1	$51.50
1	L03	0301	LAB/CHEMISTRY	84153	1	$92.25
1	C01	0331	CHEMOTHER/INJ	96402	1	$168.50
1	D01	0636	DRUGS/DETAIL CODE	J9217	1	$507.77
2	D02	0250	PHARMACY		6	$306.04
2	S01	0270	MED-SUR SUPPLIES		1	$11.25
2	S02	0270	MED-SUR SUPPLIES	A6257	1	$0.42
2	L04	0305	LAB/HEMATOLOGY	85025	1	$107.00
2	C02	0335	CHEMOTHERP-IV	96413	1	$495.00
2	E13	0510	CLINIC	99213	1	$136.50
2	D03	0636	DRUGS/DETAIL CODE	J2405	8	$50.65
2	D04	0636	DRUGS/DETAIL CODE	J9305	100	$10,714.56
3	S03	0270	MED-SUR SUPPLIES		1	$11.25
3	S02	0270	MED-SUR SUPPLIES	A6257	1	$0.42
3	E13	0510	CLINIC	99213	1	$136.50

Table 5.7a. Services

Services		
Account number	**CDM item**	**Units**
1	L01	1
1	L02	1
1	L03	1
1	C01	1
1	D01	1
2	D02	6
2	S01	1
2	S02	1
2	L04	1
2	C02	1
2	E13	1
2	D03	8

2	D04	100
3	S03	1
3	S02	1
3	E13	1

Table 5.7b. Charge description master

Charge description master			
CDM item	**Revenue code**	**HCPCS code**	**Unit charge**
C01	0331	96402	$168.50
C02	0335	96413	$495.00
D01	0636	J9217	$507.77
D02	0250		$51.01
D03	0636	J 2405	$6.33
D04	0636	J 9305	$107.15
E13	0510	99213	$136.50
L01	0300	36415	$55.75
L02	0300	84066	$51.50
L03	0301	84153	$92.25
L04	0305	85025	$107.00
S01	0270		$11.25
S02	0270	A6257	$0.42
S03	0270		$11.25

Table 5.7c. Revenue codes

Revenue code	Revenue code definition
0250	PHARMACY
0270	MED-SUR SUPPLIES
0300	LABORATORY OR LAB
0301	LAB/CHEMISTRY
0305	LAB/HEMATOLOGY
0331	CHEMOTHER/INJ
0335	CHEMOTHERP-IV
0510	CLINIC
0636	DRUGS/DETAIL CODE

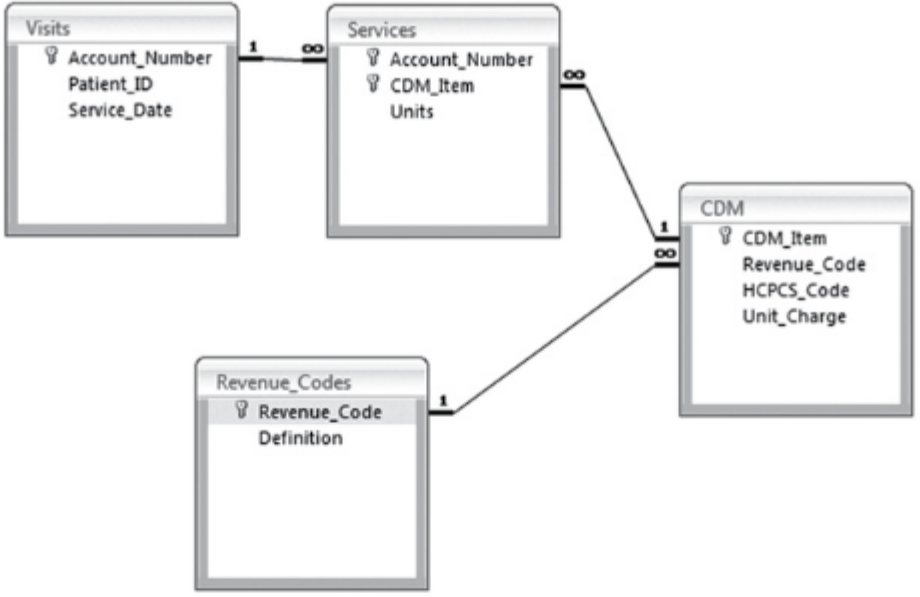


Figure 5.2. Relationships between normalized tables

▶ Expand/collapse Extended Description

Used with permission from Microsoft.

Check Your Understanding 5.1

Answer the following questions.

1. Which database type stores both the data element and the method for its use?
   a. Relational database
   b. Object-oriented database
   c. Hierarchical database
   d. Flat file database
2. A primary key
   a. Uniquely identifies a record
   b. Contains only numerical values
   c. May be repeated across rows
   d. Represents the relationship between two fields
3. Comma-Seperated Values (CSV) is:
   a. information stored in columns
   b. stores information separated by commas
   c. Allows information to be shared in a simplified format.
   d. Both B and C

Indicate whether each of the following is true or false.

4. The fields in a database represent the rows of the database.
5. The records in a database are the rows of data in the database.
6. Cardinality represents the number of elements between two tables.

Data Dictionary

A **data dictionary** is a tool that provides metadata, or data about data. For example, the data element First Name might include metadata such as the field length, that it is an alphanumeric field, and that the field is required. A key focus of a data dictionary is "to support and adopt more consistent use of data elements and terminology to improve the use of data in reporting. A data dictionary provides standardization to promote clearer understanding; promotes consistent and efficient use of information; and supports better data management" (AHIMA 2017; Dooling et al. 2014, 7).

A data dictionary should document the following attributes (or metadata) for the fields in a data table at a minimum:

- Field name
- Table name
- Description of the field
- Data type (numerical or text, field length)
- Data frequency (required field or not)
- Primary or foreign key
- Valid values
- Data source
- Field creation date
- Field termination date
- Update frequency

The data dictionary in table 5.8 documents the contents of the CDM table found in the normalization example. The details included in this table allow a user with little or no knowledge of the contents of the table to understand the fields present and their various roles.

Table 5.8. Data dictionary

Field Name	Table	Description	Data Type	Field Length	Data Frequency	Key?	Valid Values	Data Source	Field Creation Date	T
CDM item	CDM	CDM item code	Text	20	Required	Primary	Alpha-numeric	Finance	01/01/1900	
Revenue code	CDM	UB-04 revenue code	Text	4	Required	Foreign	Valid revenue code	HIM	01/01/1900	
HCPCS code	CDM	CPT/HCPCS Level II	Text	5		Foreign	Valid HCPCS	HIM	01/01/1900	
Unit charge	CDM	Charge per unit	Currency	10	Required		>= $0.00	Finance	01/01/1900	

Structured Query Language (SQL)

Structured Query Language (SQL) is the programming language that is used to manipulate data in a relational database. SQL is sometimes pronounced "sequel." SQL fulfills many roles in a DBMS:

- SQL is an interactive query language. Users type SQL commands into an interactive SQL program to retrieve data and display it on the screen, providing a convenient,

easy-to-use tool for ad hoc database queries.

- SQL is a database programming language. Programmers embed SQL commands into their application programs to access the data in a database. Both user-written programs and database utility programs (such as report writers and data entry tools) use this technique for database access.
- SQL is a database administration language. The database administrator responsible for managing a minicomputer or mainframe database uses SQL to define the database structure and control access to the stored data.
- SQL is a client/server language. Personal computer programs use SQL to communicate over a network with database servers that store shared data. This client/server architecture has become very popular for enterprise-class applications.
- SQL is an Internet data access language. Internet web servers that interact with corporate data and Internet application servers all use SQL as a standard language for accessing corporate databases.
- SQL is a distributed database language. Distributed DBMSs use SQL to help distribute data across many connected computer systems. The DBMS software on each system uses SQL to communicate with the other systems, sending requests for data access.
- SQL is a database gateway language. In a computer network with a mix of different DBMS products, SQL is often used in a gateway that allows one brand of DBMS to communicate with another brand (Groff et al. 2010, 6).

The role of SQL as an interactive query language is described in more detail in the following section. The other roles of SQL are beyond the scope of this text, but health information management (HIM) professionals should be aware that SQL is far more than just a tool used to extract data from a relational database.

SQL commands that may be used to retrieve data follow a basic structure that allows nonprogrammers to understand and write queries. For example, a SQL command to collect all of the visits for a particular patient from a visits table is

SELECT * FROM visits WHERE patient_id = '987ZYX'

The results of the query are the two records displayed in table 5.9. This is called a **select query**. The purpose of a select query is to pull records that meet a certain criterion from a particular table or combination of tables. The words SELECT, FROM, and WHERE are intentionally capitalized in the query to indicate that they are keywords or instructions in the SQL command.

In this case, the two visit records for patient 987ZYX were selected from the three records found in the visits table. The query may be modified to present the records in a particular order. For instance, the records in the previous example may be sorted by descending service date by revising the query to include an ORDER BY statement:

```
SELECT Visits.Account_Number, Visits.Patient_ID, Visits.Service_Date
FROM Visits
WHERE Visits.Patient_ID="987ZYX"
ORDER BY Visits.Service_Date DESC;
```

The ORDER BY clause states that the results of the query should be sorted by descending service date. The results of this query are presented in table 5.10.

Table 5.9. Select query

Field ermination Date	Update Frequency
	N/A
	Annual
	Annual
	Annual

Account_Number	Patient_ID	Service_Date
2	987ZYX	01/04/2011
3	987ZYX	02/06/2011

Table 5.10. Select query example results

Account_Number	Patient_ID	Service_Date
3	987ZYX	02/06/2011
2	987ZYX	01/04/2011

A select query may be used to combine data from two or more tables by defining a **join**. For example, a SQL query to select all of the patients that received the medical supply with the CDM item code S02 would require the services table be joined to the visits table. The relationship between these tables is found in the ERD displayed in figure 5.2. The tables must be combined or joined based on the Account_Number field. When more than one table is included in a query, the fields are referenced by both the table and the field name concatenated together and delimited by a period. For example, when referring to the field Account_Number in the visits table, the user would reference the field as Visits.Account_Number. The following select query will find the patients who received CDM item S02:

```
SELECT Visits.Account_Number, Visits.Patient_ID, Visits.Service_Date,
Services.CDM_Item, Services.Units
FROM Visits INNER JOIN Services ON Visits.Account_Number =
Services.Account_Number
WHERE Services.CDM_Item="S02";
```

In this query, the SELECT statement includes the fields that are desired; the FROM statement includes the two tables (Visits and Services) as well as how those two tables should be related (Account_Number). Finally, the WHERE statement indicates only rows where the CDM_Item is equal to S02 in the services table. The results of this query are presented in table 5.11.

A SQL command may also be used to summarize data. In the above example, a count of the number of visits for each patient may be generated with the following query:

```
SELECT Visits.Patient_ID, Count(Visits.Service_Date) AS CountOfService_Date
FROM Visits
GROUP BY Visits.Patient_ID;
```


Table 5.11. Join query example results

Account_Number	Patient_ID	Service_Date	CDM_Item	Units
2	987ZYX	1/4/2011	S02	1
3	987ZYX	2/6/2011	S02	1

Table 5.12. Count of service dates

Patient_ID	CountOfService_Date
123ABC	1

987ZYX	2

This query includes the familiar SELECT and FROM keywords, but notice that there is now a GROUP BY statement. When a query is used to summarize data the summary statistic and the field that designates the group to be summarized must be designated. In the example, the SELECT statement includes a command that creates a count of the number of service dates, Count(Visits,Service_Date), and stores that result in a new field called CountOfService_Date. The GROUP BY portion of the query states that the counts should be presented by Patient_ID. The result of this query is presented in table 5.12.

Check Your Understanding 5.2

Answer the following questions.

1. SQL is an acronym for
   a. Sequential Query Language
   b. Structured Query Language
   c. Selective Question Language
   d. Sequential Question Language
2. SQL is used to manipulate
   a. Data in a relational database
   b. Data in a flat file database
   c. A data warehouse
   d. Unstructured data
3. What keyword is missing from this SQL statement: SELECT name, birthdate WHERE month(birthdate) = 12
   a. BY
   b. FROM
   c. LIST
   d. GROUP

4. Which of the following SQL keywords results in an aggregation of data?
   a. FROM
   b. SELECT
   c. GROUP BY
   d. WHERE
5. A field reference of visits.patient_id refers to:
   a. A field called visits in the patient_id table
   b. A table called patient_id in the visits database
   c. A field called patient_id in the visits table
   d. A row with a value of 'patient_id' in a field called visits

Data Modeling

A **data model** is a representation of the data to be stored in a database and the relationships between the tables and data fields. Prior to the design of the database, the business process to be supported by the database must be modeled. This can be accomplished through the use of process and data flow diagrams. The components of a database, discussed previously in this chapter, are designed and assembled in a way that supports the business process requirements. Data modeling can be carried out using either

an object-oriented or an entity relationship approach. The entity relationship approach will be outlined here.

There are two important outputs to the data modeling process. The first output is the ERD, a graphical representation of the tables in the database and their relationships. The second output is the data dictionary. The data model serves as the basis for the database structure and design as documented by the ERD and data dictionary.

The process of data modeling includes three steps: the conceptual model, the logical model, and the physical data model. The conceptual model includes a mapping of the business requirements for the database using nontechnical terms that end users can understand. A database that does not meet the needs of the end users is of no value, so they must be involved in the design process at the earliest stage. The conceptual data model is independent of the type of database that will ultimately be used to store the data. The conceptual data model may be mapped using a context-level **data flow diagram**, which maps out the database's boundary and scope. Figure 5.3 shows a context-level data flow diagram for the claims database example used earlier in this chapter. Notice that the diagram does not include any field or table information. It simply displays the general categories of data and their roles in the business process. The details regarding the actual database structure are fleshed out in the logical modeling step.

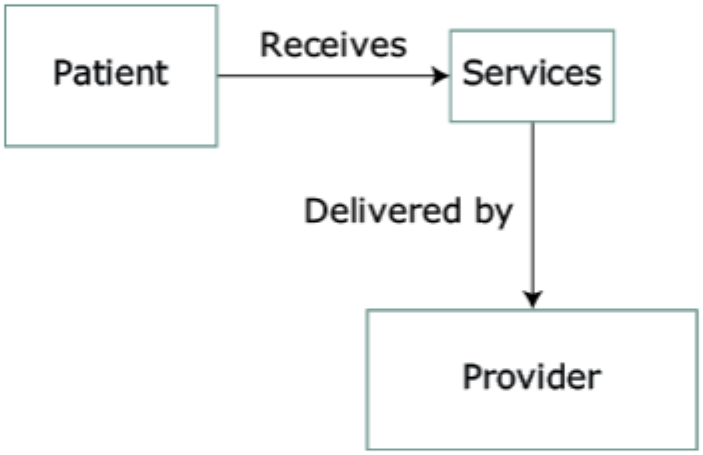


Figure 5.3. Context-level data flow diagram

▶ Expand/collapse Extended Description

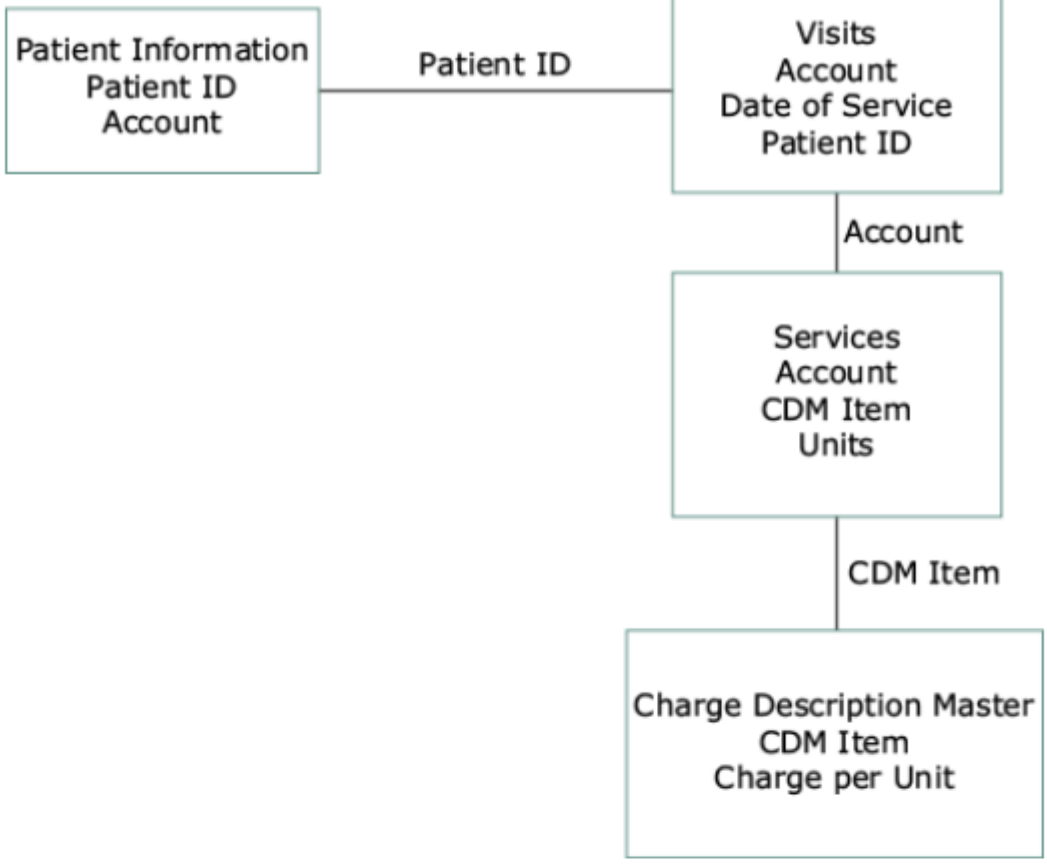


Figure 5.4. Diagram 0

▶ Expand/collapse Extended Description

The logical model is the next step in database modeling. During this phase the tables start to take shape. The fields are not defined, but the basic contents of the tables and how they relate are defined. In the claims data example, the tables are

- Patient information
- Visit for each patient
- Service(s) provided to patient
- Charge for service(s)

page 132

The relationships between the tables can be mapped out in a **Diagram 0** data flow diagram. A Diagram 0 data flow diagram expands on the context diagram and adds details regarding the tables and their relationships. Figure 5.4 displays a Diagram 0 for the claims database example.

The final step in the database modeling process is to determine the physical model. This includes mapping out the tables, keys, and relationships using an ERD as well as determining the DBMS and hardware to house the database. The ERD for this example is displayed in figure 5.2. At this point in the modeling process a database type may be selected. Because this database requires multiple tables that might need to be linked together for reporting purposes, a relational database is the best choice for this business process. There are a number of DBMSs that support relational database structures. Examples include the following:

page 133

- Microsoft Access

- Microsoft SQL Server
- Oracle
- MySQL
- PostgreSQL

The choice of specific DBMS software should be based on the size and scope of the database as well as the budget for the project.

The physical model also includes a specification of the computer hardware that will be used to house the database. The amount of storage or disk space required for the database should be determined prior to selecting any hardware platform. Estimating the space required for the database currently and for the foreseeable future will help determine the appropriate hardware and software to select for implementation of the database.

Check Your Understanding 5.3

Answer the following questions.

1. Which phase of data modeling includes a description of the business process?
   a. Physical model
   b. Logical model
   c. Conceptual model
   d. Final model
2. Which phase of data modeling describes the relationship between the data tables?
   a. Physical model
   b. Logical model
   c. Conceptual model
   d. Final model
3. What is one of the two outputs from a data modeling process?
   a. ERD
   b. SQL
   c. NDC
   d. DBMS
4. Which portion of the data modeling process maps out the relationships between the tables?
   a. Context-level data flow diagram
   b. Conceptual model
   c. Diagram 0
   d. Diagram 2

5. The choice of which DBMS to use depends on:
   a. Physical data model
   b. Conceptual data model
   c. Size and scope of the database
   d. Availability of physical space

Data Warehouses and Data Marts

Arraying data into databases for storage and retrieval is the first step in using data for decision-making. The case study in this chapter was a very simple database designed to

hold information about the services provided to a patient. Clinical and operational databases can be found in various departments throughout a healthcare facility, including

- Patient accounting
- Clinical departments
- Pharmacy
- Clinical trials
- Marketing
- Quality measurement
- Finance
- Patient experience

Combining these data sources into a **data warehouse** is an effective method of making the data accessible to a wider audience of users. A data warehouse must be designed in a way that maintains sufficient detail so that the data are useful for research and business intelligence (BI) activities, such as analytic purposes, and related in a way that supports real-time access to the data by users. Data warehouses are snapshots of a variety of databases found throughout a company that are combined for the purpose of reporting and analysis. The data warehouse does not include the "live" or transaction data that are used for business operations, but instead is populated by periodic downloads from the live database. A **clinical data warehouse** is a database that facilitates accessing data from many databases in one query for reporting (AHIMA 2017). The frequency of refreshing the clinical data warehouse depends on the volatility of the underlying data (how often it changes) and the need for timely reporting.

The key requirements for a data warehouse and business intelligence system (DW/BI) include:

1. The DW/BI system must make information easily accessible.
2. The DW/BI system must present information consistently.
3. The DW/BI system must adapt to change.
4. The DW/BI system must present information in a timely way.
5. The DW/BI system must be a secure bastion that protects information assets.
6. The DW/BI system must serve as the authoritative and trustworthy foundation for improved decision-making.
7. The business community must accept the DW/BI system to deem it successful (Kimball and Ross 2013, 3).

Although these requirements were not specifically written about clinical or healthcare data warehouses, all of the requirements still apply. The last requirement, acceptance by the business community, is key. If the data warehouse is not accepted as the source of truth for data used to drive operations, then there is no reason for it to exist. Data warehouses can often become unwieldy and difficult to navigate for end users. To avoid this issue, departments often construct **data marts**, which are well-organized, user-centered, searchable databases containing information drawn from a data warehouse to meet the users' specific needs (such as specific data in the database; AHIMA 2017).

Data marts offer a convenient method of accessing the most relevant data elements for a particular use case or department, but they represent a challenge in terms of data and information governance. As such, they are not generally accepted in large organizations as a good solution for long-term data needs.

Data warehouses and data marts are both methods used to bring complex data from multiple sources to a common database for end user analysis. The data are moved from the source to the data warehouse via a process that is called **extract, transform, and load (ETL)**. Extracting data is the process of copying the data into the ETL system for manipulation (Kimball and Ross 2013, 19). Once the data are extracted, they are essentially part of the data warehouse. Because data may come from multiple systems, the transform step ensures the data are compatible. Any data mapping, correction, or translation occurs during this step. The final step in moving data from the source to the data warehouse is the load step. This is the step when the data are physically loaded into tables that make up the data warehouse. Once the ETL process is mapped out and built, it is often scheduled to occur at regular intervals, such as nightly or weekly. The data in the data warehouse are then essentially a snapshot in time of the data in the source systems at the time of the ETL process.

Decision support databases are a common example of a data warehouse. These databases are found in many healthcare entities and may include claims data, financial data, and quality data combined in one database to support both internal and external reporting. One of the unique challenges in combining clinical data is that any release of those data may be subject to HIPAA regulations if it includes any of the protected health information fields such as dates or Social Security numbers (45 CFR Parts 160, 162, and 164). Any efforts to combine data into a clinical data warehouse should include the involvement of a compliance officer or other professional with knowledge of the rules and regulations surrounding deidentification of data and the release of identifiable data.

Much of the time and effort invested in compiling a clinical data warehouse is used to clean and combine the data. Although clinical data are singularly focused on healthcare and patient information, the coding systems might or might not be consistent. For instance, pharmacy databases might include the National Drug Code (NDC) or may only include an inventory or CDM code that only has meaning inside a particular facility or system of facilities. The mapping between code sets and standardization of coding systems is an age-old issue in healthcare data. The concept of common **ontology** or a system of categories as a method of combining data that can be categorized by clinical classification systems such as ICD and CPT codes as well as those categorized by clinical nomenclatures such as Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) has emerged. An effective ontology is "a theory of those higher-level categories which structure the biomedical domain, the representation of which needs to be both unified and coherent if it is to serve as the basis for terminologies and coding systems that have the requisite degree and type of interoperability" (Smith and Ceusters 2007). A properly designed ontology could allow the automated combining of data from many sources. Such a system would speed the design and delivery of new clinical data warehouses.

Case Study: Stanford Medicine Research Data Repository (STARR)

The Stanford School of Medicine has an extensive clinical data warehouse called STARR (Stanford Medicine Research Data Repository), formally named STRIDE (Stanford Translational Research Integrated Database Environment) that stores data that supports clinical and translational research (Stanford Center for Clinical Informatics 2022). STARR includes data from both hospitals affiliated with Stanford. The hospitals use different

billing systems and yet the warehouse is designed to combine the data sets across the facilities. The STARR warehouse comprises over 1.5 million patients with over 6 million clinical data points gathered over the years. Clinical data warehouses like Stanford's STARR (STRIDE) are becoming widespread as the technology advances to support the storage and access of huge volumes of data to help with clinical research.

One project that shows tremendous promise in facilitating the combining and sharing of clinical data is a study funded by the National Institutes of Health called Informatics for Integrating Biology and the Bedside (i2b2; Murphy et al. 2010). The mission of the project is to provide researchers with a framework to combine medical record and clinical research data. The data are stored in a deidentified format so that researchers can analyze the data to develop hypotheses and perform initial analyses to investigate the likelihood for the availability of patient cohorts for clinical trials. i2b2 facilitates the sharing of large amounts of standardized clinical data and yet protects the identification of individual patients. It can be implemented using either an Oracle or a Microsoft SQL Server database.

Additional information on this research implementation (Wagholikar et al. 2018) and i2b2 can be located at Stanford's research website (Stanford Center for Clinical Informatics 2022).

Chapter 5 Review Discussions and Exercises

Answer the following questions:

1. DBMS is an acronym for
   a. Database manipulation standards
   b. Database management standards
   c. Database management system
   d. Database manipulation system
2. A flat file database is a
   a. Database that only has simple data tables
   b. Database with only one table
   c. Database that contains records with a large number of fields
   d. Database manipulation system
3. An entity relationship diagram is used to display
   a. The update schedule for data
   b. The flow of data to reports
   c. The relationships between tables in a database
   d. The number of records in each table
4. A one-to-one (1:1) relationship describes which aspect of the relationship between two tables?
   a. The number of matching fields
   b. The cardinality
   c. The number of key fields
   d. The number of rows in each table
5. Patient and Social Security number is an example of what kind of relationship?
   a. 1:1
   b. 1:N

c. N:1
d. N:N

6. Which normal form eliminates repeating groups of values?
   a. First
   b. Second
   c. Third
   d. Fourth

7. Which of the following attributes are not included in a data dictionary?
   a. Valid values
   b. Data type
   c. Update frequency
   d. Size of table
8. Which of the following SQL statements will select all columns and rows from the table visits?
   a. SELECT patient_id FROM visits
   b. SELECT * FROM visits WHERE service_date = #1/1/2010#
   c. SELECT * FROM visits
   d. SELECT account, count(patient_id) FROM visits
9. The field reference cpt.code in a SQL query refers to
   a. The database named cpt.code
   b. The field code in the table named cpt
   c. The field cpt in the table named code
   d. The table named code found in the database named cpt
10. Which phase of data modeling includes the selection of the database type?
    a. Physical model
    b. Logical model
    c. Conceptual model
    d. Final model

Discussion

11. What are the benefits of utilizing data warehouses and data marts?
12. Advantages of using SQL?
13. What are some type of questions a researcher might be able to answer using STARR data? Are there any concerns with using this data? If yes, what is one example?

REFERENCES

45 CFR Parts 160, 162, and 164. HIPAA Transactions and Code Sets, Security Rule, Privacy Rule, Breach Notification for Unsecured Protected Health Information. Department of Health and Human Services.

AHIMA. 2017. ***Pocket Glossary of Health Information Management and Technology***. 5th ed. Chicago: AHIMA Press.

Codd, E. F. 1970. A relational model of data for large shared data banks. Edited by P. Baxendate. ***Communications of the ACM*** 13(6):377–387.

Dooling, J., P. Goyal, L. Hyde, L. Kadlec, and S. White. 2014. ***Health Data Analysis Toolkit.*** Chicago: AHIMA.

Groff, J., P. Weinberg, and A. Oppel. 2010. ***SQL: The Complete Reference***. 3rd ed. New York: McGraw-Hill/Osborne.

Kimball, R., and M. Ross. 2013. ***The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling.*** 3rd ed. Indianapolis, IN: Wiley.

Microsoft. 2021. "How to convert an Access database to SQL Server." Microsoft Support. https://support.microsoft.com/en-us/office/migrate-an-access-database-to-sql-server-7bac0438-498a-4f53-b17b-cc22fc42c979.

Microsoft. 2022. "Description of database normalization basics." Microsoft Support. https://docs.microsoft.com/en-us/office/troubleshoot/access/database-normalization-description.

Murphy, S., G. Weber, M. Mendis, V. Gainer, H. Chueh, S. Churchill, and I. Kohane. 2010. Serving the enterprise and beyond with Informatics for Integrating Biology and the Bedside (i2b2). ***Journal of the American Medical Informatics Association*** 17(2):124–130.

Smith, B., and W. Ceusters. 2007. Ontology as the Core Discipline of Biomedical Informatics: Legacies of the Past and Recommendations for the Future Direction of Research. In ***Computing, Philosophy, and Cognitive Science—The Nexus and the Liminal***, edited by G.D. Crnkovic and S. Stuart, 104–122. Cambridge, UK: Cambridge Scholars Press.

Stanford Center for Clinical Informatics. 2022. Stanford Medicine Research IT Technology & Digital Solutions. https://clinicalinformatics.stanford.edu/projects/cdw.html.

Wagholikar, K. B., P. Dessai, J. Sanz, M. E. Mendis, D. S. Bell, and S. N. Murphy. 2018. Implementation of Informatics for Integrating Biology and the Bedside (i2b2) platform as Docker containers. ***BMC Medical Informatics and Decision Making*** 18(1):66.

ADDITIONAL RESOURCES

Inmon, W. H. 2005. ***Building the Data Warehouse***, 4th ed. New York: Wiley Publishing.

Knaflic, Cole Nussbaumer, 2019 Storytelling with Data: Let's Practice, Wiley ISBN-13: 978-1119621492

Knaflic, Cole Nussbaumer, 2015 Storytelling with Data: A Data Visualization Guide for Business Professionals, Wiley ISBN-13: 978-1119002253

Knowles, Chad.2022 SQL for Data Analytics: perform fast and efficient data nalysis with the power of SQL, Packt Publishing, ICBN-13-979-8354008353

Chapter

6

page 141

Healthcare Informatics and Decision Support

Learning Objectives

- Interpret the use of decision support systems (DSSs) in healthcare settings
- Compare and contrast the potential and limitations of DSSs in research
- Use data visualization processes to facilitate decision-making
- Analyze the trends that demonstrate quality, safety, and effectiveness of patient care using statistical analysis of healthcare data
- Evaluate administrative reports using department software tools

KEY TERMS

Clinical analytics
Clinical decision support system (CDSS)
Computer-assisted coding (CAC)
Computerized provider order entry (CPOE)
C-suite
Data analytics
Decision support system (DSS)
Electronic patient portals
HL7 FHIR®
Knowledge management (KM)
Natural language processing (NLP)
Unintended consequences

page 142

Decisions are a part of daily living. Through repetition, rules, and regulations for safety and well-being, along with ongoing input from our environment, humans automatically respond to situations and make instant decisions. Ignoring warnings can result in being harmed. For example, if the gas light comes on while driving a car, that tells the driver to

refuel. A red oil light might mean severe engine damage could occur if the triggering cause is not addressed immediately. Traffic lights flash warnings for when to walk, when to stop walking, or when to go (green), slow down (yellow), and stop (red), which prevent pedestrians and motorists from being harmed. Additionally, laws and regulations must be followed to avoid penalties. In a complex, changing environment, decisions must be adapted to environmental changes, and those environmental adaptations could affect future decisions.

In healthcare, which system is complex and constantly adapting to external and internal forces, making decisions may be instantaneous, such as the decision to perform cardiopulmonary resuscitation (CPR) that relies on automatic responses and trusted clinical protocols, or might involve long-term healthcare solutions with multiple factors that need human interaction and electronic data interpretation. Furthermore, the role of electronic resources in guiding the human decision-making process is evolving with the development of new applications that enable improvements in patient safety and in the provisioning of quality healthcare.

In this chapter, we examine how the healthcare environment and healthcare organizations are affected by the administrative and clinical electronic **decision support systems (DSSs)**.

In general, a DSS is a computer-based system: It gathers data from various sources and assists in providing structure to the data by "using various analytical models and visual tools in order to facilitate and improve the ultimate outcome in decision-making tasks associated with nonroutine and nonrepetitive problems" (AHIMA 2017). In an administrative decision support system, it supports organizational decisions using data sources that could include reimbursement, utilization of services, and aggregate patient sociological data. Similarly, a **clinical decision support system (CDSS)** can include alerts in the electronic health record (EHR), such as an allergy to medications, reminders for preventive healthcare services, and links for providers to find references or order sets to assists clinicians in making patient care decisions (Sandefer 2020). Here, the order sets are part of computerized provider order entry, which appropriately groups medical orders to increase the efficiency of ordering through reducing variations in the ordering process and encouraging compliance with the best practices (Payne et al. 2003 and Zhang et al. 2018).

The increased use of CDSSs has improved the quality of patient care. For example,the vast functions of CDSSs can include, but not limited to, alarm systems, diagnostics, prescription, and drug control; and these functions can manifest as computerized reminders, alters, order sets, computerized guidelines, and patient data reports (Kabene 2010 and Sutton et al. 2020).

To meet meaningful use, the "Five Rights" concept is a best framework to follow to implement CDSS options in practice:

1. Right information
2. To the right person
3. In the right intervention format
4. Through the right channel
5. At the right time in the workflow (Osheroff et al. 2012 and CDC 2022).

The challenge to a healthcare organization is to analyze any computerized system to determine how these five components can be provided. Therefore, each organization needs to define the architectural components of their system and determine the interoperability of their software programs to provide the "five rights" for their CDSS. At the same time, interoperability requires a consensus on the standards that organizations

will use to exchange electronic data with other CDSSs. The OpenCDS consortium is a multi-institutional, collaborative effort to develop open-source, standard CDSS tools and resources that support HL7 FHIR (OpenCDS 2022). **HL7 FHIR** (Fast Healthcare Interoperability Resources) is one standard framework created by the Health Level Seven International organization (HL7) for the exchange, sharing, integration, and retrieval of electronic health information to support clinical practice and the health services management (HL7 2022). Many clinical and administrative problems can be solved with FHIR components in many contexts, such as phone apps, the cloud, and EHR-based data sharing.

Knowledge Management

This section presents **Knowledge Management (KM)** systems. It is a dynamic way to transform discrete electronic health data into easily distributed information to provide context for decision-making (O'Dell 2020). In the ever-changing healthcare system, the focus is on the ability of healthcare organizations to enhance the quality of patient care. The use of KM systems to support decisions at many levels in the healthcare organization is essential. Understanding how to extract, or "mine," the patient data from databases, data stores, and discrete electronic health documentation is key to providing feedback to improve patient care. Additionally, data can be derived from evidence-based knowledge or data extracted from data.

KM subsystem functions include the evaluation of organizational needs (at high levels or discrete department, clinical levels), processing appropriate information gathered based on the needs of an individual or the organization, mining the data, and disseminating the data as processed information. For instance, individual patients can mine their information through patient portals that allow patients to graph their lab results over time. When visiting a primary care provider, that information can be discussed for a better understanding of a persistent problem (such as repeated, unexplained urinary tract infections). A researcher may mine systems of data collected from questionnaires such as Centers for Medicare and Medicaid Services (CMS) Quality Assurance and Practice Improvement surveys of patients with end stage renal disease (ESRD) (CMS 2016). Financial administrators may wish to use these same data to analyze claims for those with ESRD.

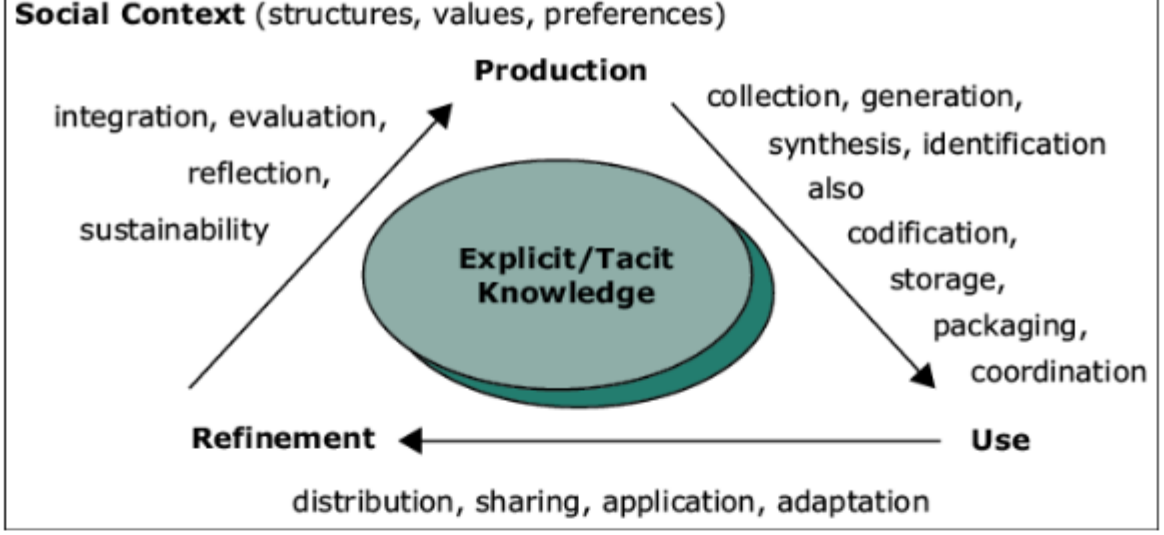


Figure 6.1. A conceptual knowledge management framework in healthcare

▶ Expand/collapse Extended Description

Source: Lau 2004, 3.

A model of KM was utilized to determine that

a. "a knowledge artifact, that is, codified knowledge is in a format that is insufficient for fully effective knowledge use by practices (e.g., digital medical records);
b. a number of interdependent processes are necessary to manage knowledge;
c. there are social and technical dimensions to these processes as a result of knowledge being tacit (e.g., that knowledge conveyed in apprenticeships relationships) and being explicit (e.g. objects such as procedure manuals);
d. action emanates from the tacit dimension of knowledge and KM processes engaged in pursuit of organization's mission". (Orzano et al. 2008)

Figure 6.1 is one conceptual model illustrating KM's inherent complexity. This includes the interdependencies of human relationships, which are essential to the success of any program. This model relies on human resources to collaborate and build systems that meet organizational performance expectations. Applied to any healthcare environment, it demonstrates an effective means to improve the quality of healthcare by accessing every employee and system in the organization. It reflects the importance of relationships, trust, and practice (social context) while expecting accessible technology and systems to generate needed information for decision-making.

Check Your Understanding 6.1

Answer the following questions.

1. Decision Support Systems provide:
   a. the source data for decision making decisions
   b. support to the electronic health record in aggregating data
   c. structure to the data by using various analytical models and visual tools
   d. are key in the meaningful use of information
2. An example of Clinical Decision Support Systems includes:
   a. allergy alerts
   b. reminders routine preventive care services

c. links to references for provider use
d. allergy alters, reminders that preventive care service are due, and links to references for provider use

3. A tool to capture and disseminate electronic information for use in decisions at multiple levels throughout the organization:
   a. Decision Support Systems
   b. Knowledge Management Systems
   c. Clinical Decision Support Systems
   d. Data Mining
4. The social issues involved with knowledge management are:
   a. Structures
   b. Values
   c. Preferences
   d. structures, values, and preferences
5. The function of an electronic decision support system (DSS) can be best described as:
   a. supporting decisions for patients
   b. supporting a healthcare organization with both administrative decisions and clinical decisions
   c. supporting OpenCDS consortium with administrative decisions
   d. Providing clinical decisions for clinicians

Administrative Decision Support Systems

Not all decisions in healthcare immediately affect the patient's safety or the quality of care delivered, but they do affect the patient's encounter with the healthcare system. For example, the process of coding a diagnosis for reimbursement affects the patient. In the most advanced systems, a code could be generated from **natural language processing (NLP)** technology systems, which are sophisticated technology that "converts human language (structured or unstructured) into data that can be translated then manipulated by computer systems" (AHIMA 2017). Through NLP, a computer would be able to understand the contents of documents, including the contextual nuances of the language within the documents, and be able to accurately extract information from the documents as well as categorize and organize the documents themselves. Leveraged by NLP, **computer-assisted coding (CAC)** has dramatically changed medical coding processes. CAC is "the process of extracting and translating dictated and then transcribed free-text data (or dictated and then computer-generated discrete data) into ICD and CPT with menu-driven prompts evaluation and management codes for billing and coding purposes" (AHIMA 2017). For example, from a patient's chart, CAC can select medical diagnosis, procedure and phrases to match the ICD and CPT codes. However, the critical analysis of a coding specialist is still needed to review the medical record to validate a final decision for medical claims. Reminders and alerts about correct coding procedures also provide clues for decision-making for the coder. Reminders to choose more specific codes, alerts that certain codes cannot be utilized for the principal diagnosis, and other reminders have increased the reliability and accuracy of coded data.

Decision support tools like NLP, reminders, and alerts are a part of the many **administrative DSSs** that affect the information processing workflow. These alerts inform patients that their claims are processed and advise healthcare organizations that the claims will be paid. Informatics professionals rely on DSSs that can assist in daily

operations. Through electronic portals, providers can send reminders to help the organization make value-based healthcare decisions. The reminders are built from the laws, rules, regulations, internal policy and procedures, and accreditation standards that healthcare organizations must follow. These tools for completing quantitative and qualitative analysis on an electronic record assist the health information specialist in completing these processes.

The **C-suite** includes executive management, such as the chief operating officer (COO) or chief information officer (CIO), as well as department managers and public health agencies. C-suite members rely on data collected and analyzed to make decisions about managing current programs, developing new programs, providing public information for patient safety improvement, and increasing overall healthcare quality. DSSs provide financial as well as clinical information for administrative decision-making. Reliance on DSS and other systems, like data warehouses and national repositories, has the potential to inform administrators of opportunities to improve efficiency and effectiveness in healthcare, which affects the quality of healthcare for patients. Electronically generated reports and dashboards provide textual information as well as visual displays of specific data from the data warehouses. The more electronically readable or structured the data are, the greater the usability for decisions. Administrators will rely on employees who can mine and statistically interpret the metadata, databases, and repositories that are rich informatics tools. This use of statistical analysis of data to make business decisions is called **data analytics** (White 2020). Similarly, **clinical analytics** is mining discrete patient healthcare data such as laboratory results, medication, genetics, or population health data to make clinical decisions or to aid in translating data for research or further healthcare treatment.

The information available to the public provides consumers with the material necessary to make intelligent decisions about where to seek healthcare organizations or individual clinicians who will meet their standards for care or will be covered by their healthcare insurers. Websites, such as HealthGrades (http://www.healthgrades.com/), the Leapfrog Group (http://www.leapfroggroup.org/), and Hospital Compare (https://www.medicare.gov/care-compare/) provide consumers with comparable information on hospitals and physicians. Social media websites, such as Facebook and Twitter, provide more direct person-to-person evaluations of healthcare experiences.

page 147

Electronic patient portals provide access to personal health information and may be available through the patient's healthcare organization. These portals provide patients with opportunities to view exam and treatment results and alerts for preventive care. Figure 6.2 shows an example of these types of reminders from the Veterans Health Administration Blue Button portal. Patients will need training on how to access their portals to facilitate patient engagement and make them aware of their responsibility in their own care. Determining if a patient has access to a computer, a mobile application, or other technology will be important to patient outcomes because these tools can provide alerts for the patients. Electronic communication systems must protect patients' rights to privacy that are compliant with all federal Health Insurance Portability and Accountability Act (HIPAA) regulations, state rules, and laws. Moreover, the importance of looking at organization-wide information systems as a way to provide meaningful information and enhance patient safety is imperative to ensure healthcare quality and safety.

page 148

Figure 6.2. Example of preventive care information from Veterans Health Administration HealtheVet Blue Button portal

VA Wellness Reminders

Source:	VA
Last Updated:	01 Aug 2017 @ 1329
Sorted By:	Name (Ascending)

Learn more about these Wellness Reminders by visiting My HealtheVet. Please contact your health care team with any questions about your VA Wellness Reminders.

Wellness Reminder	Due Date	Last Completed	Location
Colon Cancer Screening	01 Oct 2022	01 Oct 2012	Portland OR VAMC
Influenza Vaccine	DUE NOW	UNKNOWN	Portland OR VAMC
Pneumonia Vaccine	DUE NOW	UNKNOWN	Portland OR VAMC
Influenza Vaccine	DUE NOW	UNKNOWN	Spokane WA VAMC
Pneumonia Vaccine	DUE NOW	UNKNOWN	Spokane WA VAMC
Colon Cancer Screening	DUE NOW	UNKNOWN	Spokane WA VAMC
Colon Cancer Screening	DUE NOW	UNKNOWN	WHITE CITY VAMC
Influenza Vaccine	DUE NOW	UNKNOWN	WHITE CITY VAMC
Pneumonia Vaccine	DUE NOW	UNKNOWN	WHITE CITY VAMC

Source: Adapted from Blue Button, Veterans Health Administration, 2021. http://www.va.gov/BLUEBUTTON/docs/mhv_MHVTESTVETERAN_20170808_0937.pdf

Case Study: Patient Engagement

Patient engagement has been shown to increase compliance with medication schedules and reduce missed office visits, which improves clinical outcomes. Patient portals, mobile applications, and other electronic and web-based tools are being studied for associations with increased patient engagement. This case study describes how a patient portal system facilitated patient engagement and medical compliance, while providing an easy, free way for patients to refill medications and obtain and update copies of their medical records.

Two retired US Army veterans use VA Blue Button to maintain up-to-date copies of their medical records and communicate with their physician's office. They also use the system for online prescription refills and secure messaging to providers. The veterans especially like the ability to view, download, and print their medical information to have this documentation ready in case of a medical emergency. Another benefit of the system is that they are less likely to forget to refill their medications or go to appointments due to the system's reminders and the user-friendly interface. In their opinion, the best feature is the cost, which is free to veterans.

This case study illustrates a success story for a patient portal serving a large audience of veterans, including many retired and physically challenged individuals. It provides online access to up-to-date copies of the patients' records, which is essential in case of an emergency. Organizations wishing to create their own portal should use the VA Blue Button system as a model. System developers should study the usability and security features.

(Source: Adapted from Peterson 2014)

Check Your Understanding 6.2

Answer the following questions.

1. A technology that converts human language (structured and Unstructured) into data that can be translated and manipulated by computer systems is:
   a. Natural Language Processing
   b. Speech Recognition Software
   c. Data Transfer
   d. Decision Support Systems

2. Data analytics is:
   a. statistical analysis software programs
   b. data mining discrete data sets
   c. use of statistical analysis of data to make business decisions
   d. the quality initiatives to assure quality in data access via DSSs
3. Which data is usually not included clinical analytics:
   a. Medication
   b. Genetics
   c. Patient's payment
   d. Laboratory results
4. An example of administrative DSSs elements includes:
   a. Natural language processing
   b. Reminders and alerts about correct coding procedures
   c. Alerts that certain codes cannot be utilized for principal diagnosis
   d. All of above
5. The use of CAC has eliminated the coding specialist positions.
   a. True
   b. False

Clinical Decision Support Systems to Improve Safety and Quality of Patient Healthcare

Electronic health record (EHR) vendors and healthcare IT teams continue to build reminders and alerts that are intended to support the goals of patient safety and quality of healthcare. The health record is the primary communication tool for a healthcare team. Using CDSS tools to improve documentation in the health record is intended to improve patient safety and overall quality of healthcare. Alerting a clinician of a potential medication interaction or a nurse of a missed medication administration will contribute to the care of the patient. Additionally, a CDSS can aid a healthcare provider in making decisions about treatment that can be used for an individual. In one example, a CDSS provided patient education while evaluating the severity of head injuries has explained decision treatments to patients while potentially reducing unnecessary high-cost CT scans for minor head injuries (Melnick et al. 2017).

Regarding the quality of CDSSs, the Standards and Interoperability Framework within the Office of the National Coordinator for Health Information Technology (ONC) orchestrates input from both public and private sectors to ensure the technical standards

and specifications are in place (ONC 2019). In collaboration with Centers for Medicare and Medicaid Services (CMS), they are continuing to integrate CDSS standards with electronic quality measurement standards (eCQI 2022)). They have developed a use case for CDSS business rules and a use case for CDSS support guidance. These use cases are now part of the rule Voluntary 2015 Edition Electronic Health Record (EHR) Certification Criteria; Interoperability Updates and Regulatory Improvements (45 CFR 170 (a)(10)). They also offer a Clinical Quality Framework (CQF) that provide tools and guidance to "facilitate the harmonization of standards for CDDS and electronic clinical quality measures (eCQM) use in clinical quality improvement" (CQF 2022). Up-to-date information on clinical protocols and disease management can be incorporated into EHRs, often reflected by alert systems. CDSSs are viewed as applications that directly affect patient care and patient safety.

Disparity in Access to Information for CDSSs

One goal of a CDSS is to eliminate provider bias for care. For instance, if a provider believes that a patient will not comply with medical management based on age, race, sexual preference, or other characteristics, information about options for treatment or medications may be withheld from them. A system-wide EHR can be used to objectively determine a patient's need for services. Alerts sent to patients and providers simultaneously based on clinical protocols and electronic data extraction could reduce bias. In developing these systems, however, patient literacy needs to be considered because low literacy levels could mean that patients are not understanding their treatment plan as presented by the physician or the nurse.

Providing patients with information related to their healthcare diagnoses contributes to improved care (Lopez et al. 2011). Regulatory agencies are now recommending the use of health information technology (HIT) systems and chronic disease management to aid in addressing disparities. For example, as the LGBTQIA+ communities have become more open, it is critical for the needs of patients from these communities to be understood by the healthcare systems serving them. The National Academy of Medicine (NAM) is part of the leadership working to decrease this disparity affecting the LGBTQIA+ communities (Lowry 2020). One movement is to increase the number of choices for gender self-identification in demographic data. In 2015, the University of California in Riverside had gender choices of Male, Female, Trans Male/Trans Man, Trans Female/Trans Woman, Gender Queer/Gender Nonconforming, or Different Identity (CBS News 2015), and now the NAM calls for further steps to standardize questions used to collect sex, gender, and sexual orientation data (Lowry 2022).

Finding the best way to communicate with a patient is essential to the selection of CDSSs that meet everyone's needs for information. Not all patients will have computers, apps on their mobile devices, or other technology to respond to alerts. Healthcare organizations that seek ways to engage patients in their healthcare should develop strategies that reduce discrimination or disparities based on lack of electronic technology availability (Gold et al. 2015; Tieu et al. 2015). The body of research that examines ways to engage patients continues to grow. The Office of the National Coordinator for Health Information Technology provides ongoing, easily accessible guidance to patients and healthcare organizations (ONC 2018).

The Joint Commission (JC) encourages focused efforts in areas where significant health disparities occur (Haider and Pronovost 2011). Disparities in healthcare continue to plague healthcare delivery systems and negatively affect patient safety and quality of care. Although equity in healthcare delivery has been emphasized in national reports and guidelines from the National Academy of Medicine, Agency for Healthcare Research and Quality (AHRQ), the US Department of Health and Human Service Office of Minority

Health, the Pew Institute, and the National Library of Medicine (NLM), it is evident that this ongoing divide needs to be reduced (Serchen et al. 2021)

Many factors in health information technology can influence ongoing patient care, for example, the disparity in access to information for CDSSs noted above. . In an area where consumers are encouraged to access their EHR, it is imperative that the tools for doing this will be available. If an alert system is in place, it is only as good as the access to the system. As decision support tool access for one community is improved, the potential for increased patient safety and quality of healthcare for others improves as well.

The CDSSs need strong leadership and executive management directives to develop strong, inclusive, patient-focused systems that also influence the efficiency and effectiveness of resources. Also key to the advancement of a stellar CDSS is the inclusion of clinician champions to lead the process. However, clinicians are not the only practitioners who need to be included. Any department utilizing an electronic clinical system, such as pharmacy, radiology, or nursing, must be part of a dynamic team to lead the organization to a successful system. End users of the system that need to be involved include health information management (HIM) administrators, quality assurance administrators (and individuals within the department who manage quality assurance), and patients.

Research about Decision Support Systems

Research on the use of DSSs should be examined to understand how their implementation and uses can support or be a barrier to the effective use of DSSs. Therefore, informaticists involved in DSSs will want to explore the plethora of articles in this area. Below are a few examples of the use of decision support tools (DSS or CDSS).

Long-Term Care

CDSSs can be used in acute- or non-acute-care settings. For example, a CDSS tool in the On-Time Quality Improvement program, namely the Pressure Ulcer Prevention (On-Time PrU), sought to design a program "to leverage the knowledge of certified nursing assistant (CNA) staff and promote proactive care coordination and planning using IT" (Hudak and Sharkey 2011). At-risk residents were reported on a Trigger Summary, which identified known risk criteria such as nutrition, bowel incontinence, weight, and Foley catheter use. The On-Time PrU focused on quality improvement processes. An important part of this CDSS tool was the integration of nursing home's existing long term care IT software and the vendor software into the typical workflow of those caring for the patient (Hudak and Sharkey 2011). Further development of CDSS tools like the On-Time PrU will benefit patient care quality, safety, and satisfaction, while reducing the healthcare costs in this environment.

Mental Health Care

To increase patient involvement in their care, a community mental health clinic compared the use of an electronic DSS to set goals with patients during care planning to promote shared decision-making and greater awareness. The result showed that case managers who used the system were more satisfied than those who did not, while the client satisfaction in the process did not improve with the DSS (Woltmann et al. 2011). The result of this study illustrates that although a DSS can contribute to easing an operational function like goal setting in mental health settings, it does not necessarily have the same effect on patient satisfaction with their care. In this case, management might need to determine if the DSS is appropriate for use or engage more patient input to improve the tool.

In another mental health study, EHR-based tools were used to screen for bipolar disorder in patients with depression (Gill et al. 2012). A screening instrument from the

World Health Organization (WHO) Composite International Diagnostic Interview (CIDI) was used as a clinical decision support tool embedded into the EHR to assist clinicians. When a patient diagnosed with depression was seen at an office visit, the screening tool automatically displayed for use. They found that EHR-based tools were useful for screening and management of bipolar disorder for patients with depression in primary care offices (Gill et al. 2012).

Mobile technology might allow for greater patient engagement. As evidenced by one study, some success was noted with improved patient engagement with an automated telephonic assessment system for remote monitoring of bilingual diabetic patients with depression from low-income areas (Ramirez et al. 2016).

Genomics and Personalized Medicine

The need for CDSS in genomics and personalized medicine is growing. The use of CDSS is complicated when protocols change for treatment, and the knowledge is not uniformly communicated. The need for national and international data repositories of knowledge content is inevitable to facilitate data integration and mining. The use of standard languages, such as Health Level 7 (HL7), the Systemized Nomenclature for Medicine (SNOMED), Logical Observation Identifiers Names and Codes (LOINC), Unified Medical Language System (UMLS), and others, continues to be of importance in this area to facilitate data merge, organization, and analysis. Additionally, how the information is collected so that it can be retrieved uniformly is of concern; for example, whether the patient data is collected from self-report or standardized tools. Furthermore, current national and international informatics standards to support the national CDSSs might conflict because different frameworks may have been tailored for specific clinical conditions. Many organizations have worked collaboratively to reduce the conflicts to promote the exchange of healthcare information and support a national infrastructure. Government agencies, such as the ONC, the AHRQ, the Veterans Health Administration, and the US Department of Health and Human Services, along with many nongovernmental agencies, have been seeking to improve the secure exchange of health information to improve patient safety and reduce healthcare costs. The need for standards continues to be a high priority for all stakeholder organizations.

It is useful to visualize the use of CDSS in various treatment areas. In one retrospective study, an enterprise-level CDSS was utilized to mine more than 800 patient records for signs of sepsis (Amland and Hahn-Cover 2019). The goal was to reduce the time between diagnosis and treatment for sepsis because delayed treatment can cause fatalities. The CDSS correctly classified 92 percent of the sepsis cases (Amand and Hahn-Cover 2019). In personalized medicine (NHGRI 2022), mining techniques are being refined to examine the association between patient traits and patient response to non-small-cell lung cancer, which can inform clinicians of the most promising treatment course (Kureshi et al. 2016). To increase HIV screening, a customized electronic record screen was used with a suite of CDSS tools to help emergency department physicians obtain consent for and process HIV tests for adults (McGuire and Moore 2016). In the second year of the HIV pilot program, 97 percent of the targeted patients were being tested. In another example, personalized healthcare for Type 2 diabetics over 60 years was effectively managed utilizing a CDSS that sent diet and exercise messages to cell phones and a website (Lim et al. 2016). One agency that is driving the CDSS use is the AHRQ. They support the use of patient-centered outcomes and CDSSs by developing CDSS tools and funding research (AHRQ 2016).

Specific skill sets are important for working with researchers to translate personalized and genomic information into meaningful information. Figure 6.3 presents skills that are

identified as important. Continuing to develop skills in research will impact the HIM and HI professional's advancement in projects that relate specifically to decision support.

Computer-Assisted Coding in All Healthcare Settings

Computer-assisted coding (CAC) started as a way to help clinical coders manage their data complexity, but now it has grown into a technology solution that provide benefits across the healthcare system (Healthcare IT News, 2021). For example, NLP has created advances in extracting key information that can be used to identify possible codes. Code numbers are used as secondary data for the revenue management cycle and data mining for registries, research studies, and administrative planning purposes. And accurate coding not only influences the capture of reimbursable claims, but also affects the entire use of informatics for those secondary purposes.

CAC DSSs need to be designed so that the designated record set to be used for coding diagnoses and procedures can be clearly utilized. This calls for a system that can track when appropriate documentation is entered into the electronic system (EHR). Additional tools such as this will assist in the coding process, which lessens the need for physician queries. A CAC total package will include decision support tools of reminders and alerts that will contribute to meeting the government regulations related to the meaningful use of EHRs (Bronnert et al. 2011). Overall, this report suggests that CAC can function to support the coding process in the following ways:

- Streamline the coding workflow
- Support clinical documentation improvement programs
- Facilitate data mining
- Create problem lists for physician review and validation
- Provide Recovery Audit Contractor (RAC) audit trails (Bronnert et al. 2011)

An important focus of CAC is the reduction in compliance risk because of the KM tools embedded in the system. Built-in fraudulent billing alerts, for instance, assure managers that misrepresentation for financial purposes is reduced or eliminated. Encoding systems will provide a multitude of resources and references that provide up-to-date coding changes and appropriate application of the current and future payment and reimbursement models.

Unintended Consequences of Clinical Decision Support Systems

When change occurs, such as reengineering a workflow or introducing a new technology, one cannot anticipate all repercussions that might occur from it. For many clinicians, HI professionals, workflow managers, and patients, human–computer interactions have been smooth after implementation of an EHR, whereas others have struggled with the change process. Also, there are unanticipated and undesirable issues or problems, which are often described as **unintended consequences** of implementation (Ash et al. 2007). For example, electronic systems, and the humans who interact with them, are not error free, and new systems could inadvertently introduce new documentation errors, communication issues, and a dependence on the system that is unrealistic. One area that might not be considered is the shift in power, control, and autonomy affecting the provider–patient relationship.

Figure 6.3. Skills for HIM professionals in the post-genomic era

Current Skills to Apply

- Policy and procedure development for managing patient health information to ensure its accuracy, integrity, privacy, and security (including investigating and resolving problems that may arise in the development of phenotype and genotype databases that involve breaches of patient information)
- Knowledge of EHR database systems design and their maintenance
- Knowledge of forms design, computer input screens, and clinical documentation tools and guidelines
- Knowledge of implementing industry standards for data sharing and interoperability
- Developing standardized healthcare data sets
- Knowledge of coding and classification systems
- Experience with organizational compliance to laws, rules, and regulations for licensure and accrediting agencies
- Knowledge of maintaining a master patient index and master client index
- Developing policies and procedures for release of medical information

Skills to Develop

- Statistics for data analysis, including calculus-based probability and statistics; statistical programming languages (SAS, S-Plus, R)
- Information science
- Computer science
- Knowledge of genetics

Source: Mendoza 2010.

In a preliminary study of **computerized provider order entry (CPOE)**, unintended consequences affecting clinicians during early implementation of the EHR were reported (Ash et al. 2007). This qualitative research demonstrated the impact of the change in workflow and frustrations that impeded successful CPOE implementations; furthermore, it revealed a widespread of the unintended consequences of CPOE in five studied hospitals (Ash et al. 2007). The AHRQ, tasked with research to improve the quality of patient care, published a publicly available Unintended Consequences Guide (Jones et al. 2011). The Unintended Consequences Guide uses the framework in figure 6.4 to illustrate the variety of issues that must be considered related to unintended consequences, as well as the iterative nature of the process, which is never truly finished.

As discussed previously, alert systems, patient bias, issues of patient engagement and connectivity to information, and basic human resistance to change will continue to perplex those analyzing data into usable information to make changes in the safety and quality of healthcare. The ever-changing dynamic nature of healthcare is consistent, and unintended consequences are to be expected.

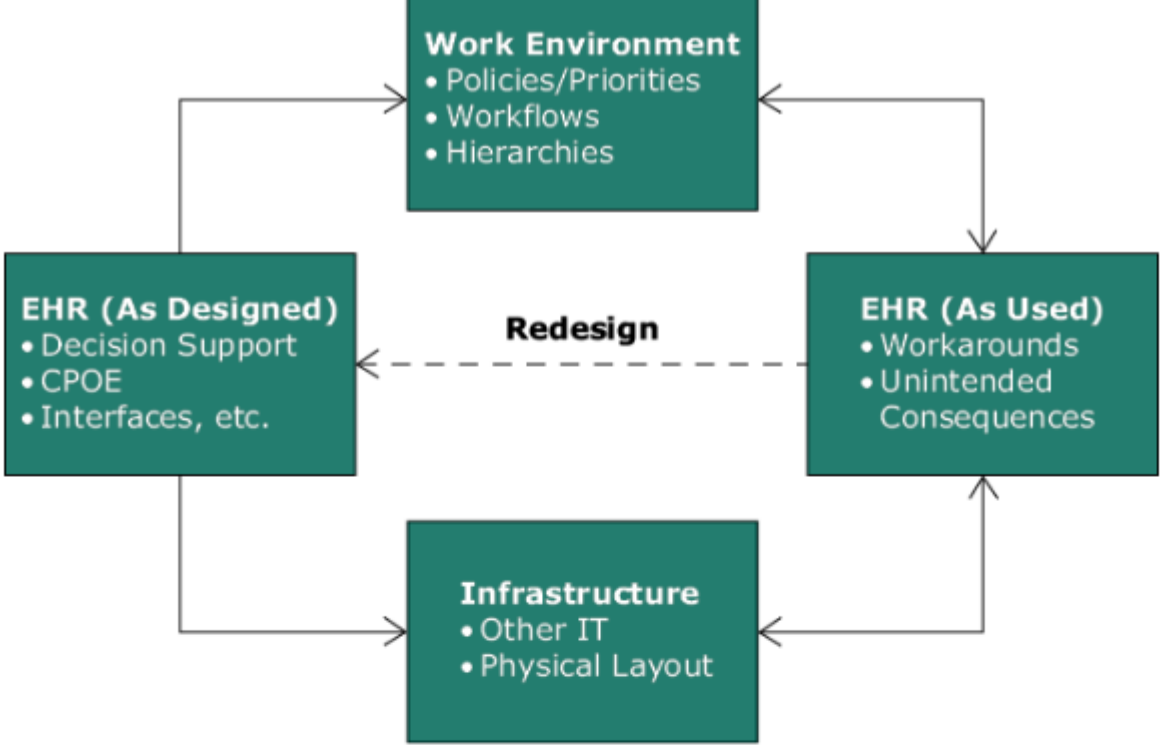


Figure 6.4. ITSA framework

▶ Expand/collapse Extended Description

Source: Jones et al. 2013.

Check Your Understanding 6.3

Answer the following questions.

1. Working around alerts:
   a. can cause legal ramifications
   b. should require physician documentation of action in response decision support
   c. is related to the definition of the legal health record
   d. all of the above
2. ONC provides ongoing guidance and is easily accessible to patients and healthcare organizations.
   a. True
   b. False
3. Potential consequences of implementing new DSSs affect:
   a. individual autonomy
   b. workflow
   c. clinical decisions
   d. all of the above
4. Disparity in healthcare cannot be improved with DSS.
   a. True
   b. False

Quality

As exemplified in the above section "Unintended Consequences of Clinical Decision Support Systems", the unintended consequences of CPOE implementation can affect the quality of healthcare; furthermore, workflow and electronic system issues can seriously and negatively affect quality of care when implementing new electronic systems. The importance of analyzing workflow prior to, during, and after implementation cannot be emphasized enough. As researchers study the implementation of a system, it is essential to consider the precise timeliness of needed information. Reflecting on current practice (what works and what does not), human relationships (trust and effective communication), and the infrastructure of systems and human resources helps in completing a successful implementation.

What continues to plague DSSs is the ability to have just-in-time information so decisions critical to patient care can be made in a timely manner. Determining who can turn alerts "off and on" and what happens when a clinician ignores an alert are other details that need to be determined. It is not only important to involve clinicians, IT professionals, quality professionals, and others in the process of determining the policies and procedures for the facilities, it is also imperative to involve legal counsel in this process. Some organizations have implemented CDSS committees to assist with setting up alerts, reviewing their effectiveness, reviewing issues, and ultimately ensuring the most effective CDSS (Kuperman et al. 2005).

One of the areas that tends to be difficult to establish for some healthcare organizations is the definition of the legal health record and the implications of ignoring or working around alerts. For example, if metadata is allowed in discovery, subpoenaed, and found admissible in a court of law, the decisions made to ignore alerts may find new legal ramifications: The court system needs to decide if and how the lack of alerts, the content of DSS, and ignoring or working around alerts have affected the quality of care of a patient. As court decisions are made related to these issues, the power and autonomy concerns will be affected. In a practice brief on defining the legal health record, AHIMA states,

> At a minimum the EHR should include documentation of the clinician's actions in response to decision support. This documentation is evidence of the clinician's decision to follow or disregard decision support. The organization should define the extent of exception documentation required (e.g., what no documentation means). (AHIMA 2011)

For HI and HIM professionals, knowledge about the legal system, the legal health record, and access to metadata are needed when DSSs and CDSSs are developed and utilized. Knowledge of laws, regulations, rules, and accreditation standards is essential to protect the safety of the patient and the healthcare organization.

The State of the Art

Although there has been significant progress in the use of CDSS, several barriers to adoption remain. A government report on CDSS design noted barriers to achieving efficient CDSS usage that included issues with inconsistent terminology, missing codes for data, low adoption of built-in order sets, poor user interface design, too many nonspecific alerts, and lagging system updates, which necessitated workarounds. The report also recommended establishing committees to validate data quality, appropriateness of order sets, rules for alerts, and screen layout (HHS 2011). Data in the CDSS must be complete and accurately coded or the alerts and reminders will not be triggered. Related report emphasized planning information workflow and the five CDSS

Rights that are right information, people, format, channel, and workflow (Wasylewicz et al. 2018).

When designing and testing CDSS systems, informatics must address potential healthcare IT safety issues that can result from poorly designed and implemented systems (National Quality Forum [NQF] 2016). For example, too many alerts will result in a clinician ignoring or disabling the alert systems (NQF 2016). CDSS content should be updated on a schedule and as needed when new information, such as drug interactions, becomes available. Quality measures must be developed and used to monitor the appropriateness of built-in orders. In turn, built-in orders set usability and must be examined with order sets updated or removed as necessary (NQF 2016). Guidance for safe CDSS use can be derived from the ONC's Safety Assurance Factors for EHR Resilience (SAFER) guidance, which advocates downtime policies, standard coding of data, clinician involvement, and regular reviews of system logic and rules (HHS 2018).

A remaining concern for HI professionals, clinicians, IT professionals, and others is the intention of those using the DSS. For example, consider a hard stop alert intended to stop a clinician from ordering two medications that will interact. Savvy clinicians might find ways to override the alert, which may be warranted by their intimate knowledge of the patient's current health situation. The art of medicine demands that the clinician have the independence to think critically about treatment. However, protecting the integrity of the legal health record by enforcing basic policies and procedures about tools such as alerts is essential. Clearly documenting who has the rights and responsibility to change alert tools further protects the facility. Conducting quality audits of alert overrides, for instance, can provide information that affects patient safety.

In the future, the CDSSs may be affected by

- Social risks—Well-trained informatics personnel need, multidisciplinary leadership requirements, CDSS policies and procedures evolving
- Technical risks—System interface interoperability, system upgrades, and maintenance must be scheduled
- Coding changes—Coded data in the system need frequent updates, legacy data must also be preserved for reporting, and allergies must be up to date across systems
- Patient identification—Challenging with data from multiple systems needing to be reconciled in CDSS (Sutton et al. 2020).

DSSs, whether clinical or administrative, call for creative individuals to continue development and manage the knowledge needed to keep current. This is time intensive and costly. Large academic hospitals, government facilities such as the Veterans Affairs Medical Centers, and large healthcare enterprises have traditionally taken the lead in development, research, and evaluation of systems. The benefits of the collective work provide smaller healthcare systems with the hope for accessing DSSs that will also enhance their administrative and patient safety goals.

Chapter 6 Review Exercises

Answer the following questions.

1. Computer-assisted coding can function to support the coding process in the following ways except:
   a. Streamline the coding workflow
   b. Create problem lists for physician review and validation
   c. Support clinical documentation improvement programs

d. All of above

2. All decisions made in the delivery of healthcare require careful study and may take time and require much study.
   a. True
   b. False
3. The use of electronic systems to support human decision making requires a change in mindset for this process to be trusted.
   a. True
   b. False

4. "Rights" must be given through the electronic health record to allow the DSS/CDSS to access the data needed for the data analysis performed by the system.
   a. True
   b. False
5. All decisions in healthcare immediately affect the patient's safety or the quality of care delivered.
   a. True
   b. False
6. Patient portals provide access to a patient's individual health information
   a. True
   b. False
7. One reason for disparity in access to patient information in CDSSs is the patient's lack of technology to use to gain access.
   a. True
   b. False
8. A Computer Assisted Coding (CAC) total package does not contribute to the meaningful use activities but does support quality coding.
   a. True
   b. False
9. A focus of Computer Assisted Coding (CAC) is to reduce compliance risk.
   a. True
   b. False

Discussion

10. Discuss the use of DSSs in healthcare settings.
11. Compare and contrast the potential and limitations of DSSs in research.
12. Identify data visualization processes to facilitate decision-making.
13. A coding manager for a large hospital has been given an annual administrative report generated by a CDSS describing coding productivity for the year. How should the coding manager evaluate the report using department software and other tools?

REFERENCES

45 CFR 170(a)(10). 2015. "Voluntary 2015 Edition Electronic Health Record (EHR) Certification Criteria; Interoperability Updates and Regulatory Improvements." https://www.law.cornell.edu/cfr/text/45/part-170/subpart-C.

Agency for Health Research and Quality (AHRQ). 2016. "***Health information technology—Best practices transforming quality, safety and efficiency.***" https://heal

thit.ahrq.gov/ahrq-funded-projects/clinical-decision-support-cds.

AHIMA. 2011. Fundamentals of the legal health record and designated record set. ***Journal of AHIMA*** 82(2). http://library.ahima.org/xpedio/groups/public/documents/ahima/bok1_048604.hcsp?dDocName=bok1_048604 .

AHIMA. 2017 ***Pocket Glossary for Health Information Management and Technology***. 5th ed. Chicago: AHIMA Press.

Amland, R. C., and K. E. Hahn-Cover. 2019. Clinical decision support for early recognition of sepsis. ***American Journal of Medical Quality*** Vol. 34(5) 494–501..https://doi.org/10.1177/1062860619873225..

Ash, J., D. Sittig, E. Poon, K. Guappone, E. Campbell, and R. Dyskstra. 2007. The extent and importance of unintended consequences related to computerized provider order entry. ***Journal of the American Medical Informatics Association*** 14(4):415–423. doi:10.1197/jamia.M2373.

Bronnert, J., B. Cassidy, S. Eichenwald-Maki, H. Eminger, J. Flanagan, M. Morsch, K. Peterson, K. Phibbs, P. Resnick, R. Schichilone, and G. Smith. 2011. ***CAC 2010-11: Industry Outlook and Resources Report.*** Chicago: AHIMA.

Campbell, R. J. 2013. The five rights of clinical decision support: CDS tools helpful for meeting meaningful use. ***Journal of AHIMA*** 84(10):42–47.

CBS News. 2015. "***California colleges will now ask students to pick from 6 genders.***" https://ucrtoday.ucr.edu/30352.

CDC. 2022. "What Is Clinical Decision Support". https://www.cdc.gov/opioids/healthcare-admins/ehr/clinical-decision-support.html.

Centers for Medicare and Medicaid Services (CMS). 2016. "End State Renal Disease (ESRF) center." https://www.cms.gov/Center/Special-Topic/End-Stage-Renal-Disease-ESRD-Center.html.

CQF. 2022. Clinical Quality Framework. https://ecqi.healthit.gov/tool/cqf.

Department of Health and Human Services, Office of the National Coordinator. 2011. "Decision support evaluation tools clinical decision support design." https://www-.healthit.gov/sites/default/files/cds/3_5_5_decision_support_evaluation_tools.pdf .

Department of Health and Human Services, Office of the National Coordinator. 2018. "Safety Assurance Factors for EHR Resilience (SAFER) guides." https://www.healthit.gov/topic/safety/safer-guides.

eCQI. 2022. Electronic Clinical Quality Improvement (eCQI) Resource Center. https://ecqi.healthit.gov/.

Gill, J., Y. X. Chen, A. Grimes, and M. Klinkman. 2012. Using health record–based tools to screen for bipolar disorder in primary care patients with depression. ***Journal of American Board of Family Medicine*** 25:283–290. doi:10.3122/jabfm.2012.03.110217.

Gold, M., M. Hossian, and A. Manqum. 2015. Consumer engagement in health IT: Distinguishing rhetoric from reality ***eGEMS*** 3(1):1190. doi: 10.13063%2F2327-9214.1190.

Haider, A., and P. Pronovost. 2011. Health information technology and the collection of race, ethnicity, and language data to reduce disparities in quality of care. ***The Joint Commission Journal on Quality and Patient Safety*** 37(10):435–436.

Healthcare IT News. 2021. "The growing benefits of computer-assisted coding technology". https://www.healthcareitnews.com/news/anz/growing-benefits-computer

-assisted-coding-technology.
HL7. 2022. "About HL7." https://www.hl7.org/about/.
HL7. 2022. "FHIR implementation standards." http://www.hl7.org/implement/standards/fhir/summary.html.
Hudak, S., and S. Sharkey. 2011. ***California Healthcare Foundation. Trendspotting: How IT Triggers Better Care in Nursing Homes***. https://www.chcf.org/wp-content/uploads/2017/12/PDF-TrendspottingTriggersBetterCareNursingHomes.pdf .
Jones, S. S., R. Koppel, M. S. Ridgely, T. E. Palen, S. Wu, and M. I. Harrison. 2011. ***Guide to Reducing Unintended Consequences of Electronic Health Records. Prepared by RAND Corporation under Contract No. HHSA290200600017I, Task Order #5.*** Rockville, MD: Agency for Healthcare Research and Quality. https://healthit.ahrq.gov/sites/default/files/docs/publication/guide-to-reducing-unintended-consequences-of-electronic-health-records.pdf .
Kabene SM. Healthcare and the Effect of Technology: Developments, Challenges and Advancements. 2010. ***Medical Information Science Reference (First edition)***.
Kuperman, G. J., R. Diamente, V. Khatu, T. Chan-Kraushar, P. Stetson, A. Boyer, and M. Cooper. 2005. ***AMIA Annual Symposium Proceedings*** 415–419.
Kureshi, N., S. S. Abidi, and C. Blouin. 2016. A predictive model for personalized therapeutic interventions in non-small-cell lung cancer . ***IEEE Journal of Biomedical Health Informatics*** 20(1):424–431. doi:10.1109/JBHI.2014.2377517.
Lau, F. 2004. Toward a conceptual knowledge management framework in health. ***Perspectives in Health Information Management*** 1:8.
Lowry, M. 2020. Understanding the Well-Being of LGBTQI+ Populations. https://www.nationalacademies.org/news/2020/07/understanding-the-well-being-of-lgbtqi-populations.
Lowry, M. 2022. NIH Should Standardize Questions Used to Collect Sex, Gender, and Sexual Orientation Data in Studies and Surveys. https://www.nationalacademies.org/news/2022/03/nih-should-standardize-questions-used-to-collect-sex-gender-and-sexual-orientation-data-in-studies-and-surveys-says-new-report .
Lim, S., S. M. Kang, K. M. Kim, J. H. Moon, S. H. Choi, H. Hwang, H. S. Jung, K. S. Park, J. O. Ryu, and H. C. Jang. 2016. Multifactorial intervention in diabetes care using real-time monitoring and tailored feedback in type 2 diabetes. ***Acta Diabetol*** 53(2):189–198. doi:10.1007/s00592-015-0754-8.
Lopez, L., A. Green, A. Tan-McGrory, R. King, and J. Betancourt. 2011. Bridging the digital divide in healthcare: The role of health information technology in addressing racial and ethnic disparities. ***The Joint Commission Journal on Quality and Patient Safety*** 37(10):437–445.
McGuire, R., and E. Moore. 2016. Using a configurable EMR and decision support tools to promote process integration for routine HIV screening in the emergency department. ***Journal of the American Medical Informatics Association*** 23(2):396–401. doi:10.1093/jamia/ocv031.

Mendoza, M. C. B. 2010. HIM and the path to personalized medicine: Opportunities in the post-genomic era. ***Journal of AHIMA*** 81(11):38–42.
Melnick, E. R., E. Hess, G. Guo, M. Breslin, K. Lopez, A. Pavlo, F. Abujarad, F. Powsner, and L. Post. 2017. Patient-centered decision support: Formative usability evaluation of integrated clinical decision support with a patient decision aid for minor

head injury in the emergency department. ***Journal of Medical Internet Research***. 19(5):e174. doi: 10.2196/jmir.7846.

Mera, M., C. Gonzalez, and D. M. Lopez. 2014. ***Towards an Intelligent Decision Support System for Public Health Surveillance—A Qualitative Analysis of Information Needs***. Amsterdam: IOS Press.

National Human Genome Research Institute (NHGRI). 2022. "Personalized Medicine." https://www.genome.gov/genetics-glossary/Personalized-Medicine.

National Library of Medicine (NLM). 2016. "Health literacy and cultural competence." https://www.nlm.nih.gov/hsrinfo/health_literacy.html.

National Quality Forum (NQF). 2016. "Identification and prioritization of HIT patient safety measures." http://www.qualityforum.org/Publications/2016/02/Identification_and_Prioritization_of_HIT_Patient_Safety_Measures.aspx .

O'Dell, R. M. 2020. Clinical Quality Management. In ***Health Information Management Concepts Principles, and Practices, 6th ed.,*** edited by P. Oachs and A. Watters. Chicago: AHIMA Press.

Office of the National Coordinator for Health Information Technology (ONC). 2019. Standards & Interoperability (S&I) Framework. https://www.healthit.gov/topic/interoperability/standards-interoperability-si-framework.

Office of the National Coordinator for Health Information Technology (ONC). 2015. "***What is eHealth Exchange?"*** https://www.healthit.gov/faq/what-ehealth-exchange.

Office of the National Coordinator for Health Information Technology (ONC). 2018. "***Patient education and engagement.***" https://www.healthit.gov/topic/patient-education-and-engagement.

OpenCDS. 2022. https://www.opencds.org/.

Orzano, J. A., C. R. McInerney, D. Scharf, A. F. Tallia, and B. F. Crabtree. 2008. A knowledge management model: Implications for enhancing quality in healthcare. ***Journal of the American Society for Information Science and Technology*** 59(3):489–505.

Osheroff, J. A, J. M. Teich, D. Levick, L. Saldana, F. T. Velasco, D. F. Sittig, K. M. Rogers, and R. A. Jenders. 2012. ***Improving Outcomes with Clinical Decision Support: An Implementer's Guide, 2nd ed***. Chicago: HIMSS.

Payne, T. H., P. J. Hoey, P. Nichol, and C. Lovis. 2003. Preparation and use of preconstructed orders, order sets, and order menus in a computerized provider order entry system. ***Journal of the American Medical Informatics Association***. 10(4):322-9. doi: 10.1197/jamia.M1090.

Peterson, H. 2014. "US Department of Veterans Affairs: The Blue Button is literally a life saver." http://www.va.gov/health/newsfeatures/2014/september/the-blue-button-is-literally-a-life-saver.asp.

Plain Language Action and Information Network. 2011. "Federal Plain Language guidelines." https://www.plainlanguage.gov/.

Ramirez, M., S. Wu, H. Jin, K. Ell, S. Gross-Schulman, L. Myerchin Sklaroff, and J. Guterman. 2016. Automated remote monitoring of depression: Acceptance among low-income patients in diabetes disease management. ***JMIR Mental Health*** 3(1):e6.

Rapporteur, L. M. 2013. ***Health Literacy: Improving Health, Health Systems, and Health Policy Around the World—Workshop Summary***. Washington, DC: National Academies Press.

Sandefer, R. 2020. Health Information Technologies. In ***Health Information Management Concepts Principles, and Practices,*** 6th ed., edited by P. Oachs and A. Watters. Chicago: AHIMA Press.

Serchen, J., R. Doherty, O. Atiq, D. Hilden, and the Health and Public Policy Committee of the American College of Physicians. 2021. A comprehensive policy framework to understand and address disparities and discrimination in health and health care: A policy paper from the American College of Physicians. Annals of Internal Medicine. 174(4):529-532. doi: 10.7326/M20-7219.

Sutton, R.T., D. Pincock, D. C. Baumgart, D. C. Sadowski, R. N. Fedorak, and K. I. Kroeker. An overview of clinical decision support systems: benefits, risks, and strategies for success. 2020. ***NPJ Digit Med***. 3:17. doi: 10.1038/s41746-020-0221-y.

Tieu, L., U. Sarkar, D. Schillinger, J. D. Ralston, N. Ratanawongsa, R. Pasick, and C. Lyles 2015. Barriers and facilitators to online portal use among patients and caregivers in a safety net health care system: A qualitative study. ***Journal of Medical Internet Research*** 3(17). doi:10.2196/jmir.4847.

Veterans Health Administration. 2016. "***Blue Button.***" http://www.va.gov/bluebutton/.

Wasylewicz, A. T. M., and A. M. J. W. Scheepers-Hoeks. Clinical Decision Support Systems. 2018 Dec 22. In: Kubben P, Dumontier M, Dekker A, editors. Fundamentals of Clinical Data Science [Internet]. Cham (CH): Springer; 2019. Chapter 11. Available from: https://www.ncbi.nlm.nih.gov/books/NBK543516/ doi: 10.1007/978-3-319-99713-1_11.

White, S. 2020. Healthcare Data Analytics. In ***Health Information Management Concepts Principles, and Practices,*** 6th ed., edited by P. Oachs and A. Watters. Chicago: AHIMA Press.

Woltmann, E., S. Wilkniss, A. Teachout, G. McHugo, and R. Drake. 2011. Trial of an electronic decision support system to facilitate shared decision making in community mental health. ***Psychiatric Services*** 62(1). http://ps.psychiatryonline.org/doi/abs/10.1176/ps.62.1.pss6201_0054.

Zhang, Y., R. Trepp, W. Wang, J. Luna, D. Vawdrey, and V. Tiase. 2018. Developing and maintaining clinical decision support using clinical knowledge and machine learning: the case of order sets. ***Journal of the American Medical Informatics Association.*** 25(11):1547-1551.doi: 10.1093/jamia/ocy099.

ADDITIONAL RESOURCES

Agency for Healthcare Research and Quality (AHRQ). 2021. ***National Healthcare Quality and Disparities Reports.*** https://www.ahrq.gov/research/findings/nhqrdr/nhqdr21/index.html.

Ahmadian, L., M. van Engen-Verheul, F. Bakhshi-Raiez, N. Peek, R. Cornet, and N. F. de Keizer. 2011. The role of standardized data and terminological systems in computerized clinical decision support systems: Literature review and survey. ***International Journal of Medical Informatics*** 80(2):81–93. doi:10.1016/j.ijmedinf.2010.11.006.

Beeler, P. E., E. J. Orav, D. L. Seger, P. C. Dykes, and D. W. Bates. 2016. Provider variation in responses to warnings: Do the same providers run stop signs repeatedly?

Journal of the American Medical Informatics Association 23(e1):e93–98. doi:10.1093/jamia/ocv117.

Berner, E. S. 2002. Ethical and legal issues in the use of clinical decision support systems. ***Journal of Health Information Management*** 16:4.

Berner, E. 2009. ***Clinical Decision Support Systems: State of the Art.*** Publication No. 09-0069-EF. Rockville, MD: ***Agency for Healthcare Research and Quality***.

Campbell, E., D. Sittig, K. Guappone, R. Dykstra, and J. Ash. 2007. Overdependence on technology: An unintended consequence of computerized provider order entry. ***AMIA Annual Symposium Proceedings*** 9:94–98.

Food and Drug Administration Safety and Innovation Act of 2012. Public Law 112-144.

Fossum, M., G. Alexander, M. Ehnfors, and A. Ehrenberg. 2011. Effects of a computerized decision support system on pressure ulcers and malnutrition in nursing homes for the elderly. ***International Journal of Medical Informatics*** 80:607–617. doi:10: 1016/j.ijmedinf.2011.06.009.

Han, Y. Y., J. A. Carcillo, S. T. Venkataraman, R. S. B. Clark, S. Watson, T. C. Nguyen, H. Bayir, and R. A. Orr. 2005. Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system. ***Pediatrics*** 116(6):1506–1512.

Kaplan, B. 2001. Evaluating informatics applications: Clinical decision support systems literature review. ***International Journal of Medical Informatics*** 64(1):15–37.

Kawamoto, K., D. Lobach, H. Willard, and G. Ginsburg. 2009. A national clinical decision support infrastructure to enable the widespread and consistent practice of genomic and personalized medicine. ***BMC Medical Informatics and Decision Making*** 9:17. doi:10:1186/1472-6947-9-17.

Kesselheim, A. S., K. Cresswell, S. Phansalkar, D. W. Bates, and A. Sheikh. 2011. Analysis and commentary: Clinical decision support systems could be modified to reduce "alert fatigue" while still minimizing the risk of litigation. ***Health Affairs*** 30:122310–122317.

Kohn, L. T., J. M. Corrigan, and M. S. Donaldson. 2001. ***Crossing the Quality Chasm: A New Health System for the 21st Century***. Washington, DC: Committee on Quality of Health Care in America, Institute of Medicine.

Learned Hand, J. ***The T.J. Hooper***, 60 F.2d 737, 740 (2d Cir. 1932).

Chapter 7

Introduction to Research

By Yanshan Wang, PhD, FAMIA

Learning Objectives

- Develop a personal vocabulary of research-related terms
- Apply the steps in the research process
- Articulate the role of the EHR in healthcare research
- Relate the principles of research design
- Examine statistical analysis tools
- Select the means of appropriate data analysis
- Choose the appropriate means for data visualization

KEY TERMS

Alternate hypothesis
ANOVA
Applied research
Basic research
Cause and effect
Clinical outcomes assessment
Control group
Correlational research
Dashboard
Data analysis
Data mining
Data retrieval
Data visualization
Dependent variable

Descriptive research
Descriptive statistics
Double-blind study
Electronic health records
Experimental research
Export
External validity
Field studies
Forecasting
Historical research
Hypotheses
Import
Independent variable
Inferential statistics
Institutional review board (IRB)
Internal validity
Metadata
Mixed methods research
Natural experiments
Nonexperimental research
Null hypotheses
Placebo
Population
Predictive analysis
Python
Qualitative research
Quantitative research
Quasi-experimental
Random sampling
REDCap
Reliability
Research
Research question
Retrospective research
Sample
Scatterplots
Structured Query Language (SQL)
t-test
Validity
Variable
Vulnerable subjects

Research is

> an inquiry process aimed at discovering new information about a subject or revising old information. Investigation or experimentation aimed at the discovery and interpretation of facts, revision of accepted theories or laws in the light of new facts, or practical application of such new or revised theories or laws; the collecting of information about a particular subject. (AHIMA 2017)

As a result of a Health Information Technology for Economic and Clinical Health (HITECH) Act modification, research is considered to be "a systemic investigation, including research development, testing and evaluation, designed to develop or contribute to generalized knowledge." **Electronic health records (EHRs)** are electronic versions of patients' past records generated from their clinical visits (Wang et al. 2018). Since the HITECH Act was signed into law in 2009, many healthcare institutions have adopted EHR systems. The primary use of EHR data is to serve clinical care and facilitate billing. As the quantity of EHR data has increased in recent decades, they have been used for research purposes, which is considered to be a secondary use of the data.

Using patient information for research purposes is quite appealing due to the capabilities of the EHR to capture or manipulate data in ways that were not previously available or that were time consuming and difficult. For example, tracking just one patient's blood pressure over several episodes of care, using paper charts, would require abstracting the readings from each of the records for the individual encounters. By contrast, with an EHR and health information exchange (HIE), the underlying databases could be queried across episodes of care and even different healthcare organizations for a specific period of time for that patient. Using the EHR, the possibilities for accessing this same information for groups of patients such as different age groups, presence of other specific diagnoses, medications and treatment provided, and so on, are almost limitless.

There are regulations specifically for using patient identifiable information for research purposes. The Health Insurance Portability and Accountability Act of 1996 (HIPAA) privacy rule requires prior patient authorization for internal use or for the disclosure of patient information for research use unless one of several exceptions are met or if approved by a privacy board or the organization's institutional review board (IRB) (45 CFR 160.514(d)). Situations that do not require patient authorization include

- Reviewing the information to plan for research to include developing the research question or hypothesis but maintaining the patient information within the covered entities that are defined in the HIPAA as health plans, healthcare clearinghouses, and healthcare providers;
- Studying those who are no longer living; and
- Using only a limited data set where data have been deidentified (removing any information that could be used to identify the individual; AHIMA 2017), also referred to as anonymized data.

The HIPAA privacy rule's minimum necessary standard requires covered entities to use the minimum personal health information necessary for the research project (45 CFR 164.502(b); 45 CFR 164.514(d)). The minimum necessary standard must be followed even when the use of the patient information is authorized or a waiver has been granted. HIPAA is the primary federal regulation that addresses the use of this information, but the Federal Policy for Protection of Human Subjects, also called the Common Rule, must be considered when support for the research has been provided by the federal government and includes entities such as the Department of Health and Human Services (HHS) and Veterans Affairs (45 CFR 46.111(a) (7)). The Common Rule is a set of regulations regarding human research and testing governed by the IRB (AHIMA 2017). Researchers must also be compliant with state laws that cover the access and use of patient data, which are, in some instances, more stringent or cover particular types of data such as for HIV or genetics.

Principles of Research Design

Consistent with the definition of research to find new information, research can be divided into two major types: basic research and applied research. **Basic research**, which can also be referred to as pure or theoretical, is done with a "focus on the development and refinement of theories" (AHIMA 2017) but not necessarily to identify a solution to the problem. **Applied research**, also referred to as practical or clinical research, "focuses on the use of scientific theories to improve actual practice" (AHIMA 2017). Moreover, it is common for the initial research on a new topic to involve exploratory studies designed to evaluate current practices for the emerging topic (Dolezel and Moczygemba 2015).

For either basic or applied research, addressing the research design before a study begins provides "structure of the study ensuring that the evidence collected will be relevant and that the evidence will unambiguously and convincingly answer the research question; the design includes a detailed plan that, in a quantitative study includes controlling variance" (AHIMA 2017). One component of the overall research design is the research approach. There are three types of research approaches used for conducting research inquiries: quantitative, qualitative, and mixed methods. Researchers must consider the purpose of the study when selecting the research approach (Layman and Watzlaf 2017). To test a hypothesis where numerical data can be collected for statistical analysis, a quantitative research approach is indicated. Conversely, qualitative studies to gather information would be appropriate to explore the perceptions of individuals or groups in a new area or regarding a shared experience.

Quantitative Research

Quantitative research aims to test a theory by defining specific, measurable study variables, collecting and analyzing data on those variables, and conducting statistical tests to examine the relationships among the variables (Creswell 2011). Quantitative research is used extensively in healthcare studies. To begin, quantitative researchers first conduct extensive literature reviews to establish the importance of the research topic, to understand what previous research has been conducted, and to define the study variables and methods based on the evidence-based work of previous researchers. Data collected for the study are generally numerical (for example, age and lab test results) and categorical (for example, gender, race, and comorbidity). Quantitative researchers seek to collect large amounts of data from many participants on the explicitly defined study variables. An example of a quantitative study would be a study on the impact of the COVID-19 pandemic on surgical neuro-oncology multidisciplinary team decision-making in which a survey was used to collect data from neurosurgical units (Price et al. 2020). This study conducted quantitative analysis to investigate whether COVID-19 had caused patients who required surgery for malignant brain tumors to receive different treatment.

Qualitative Research

Qualitative research involves investigating nonnumerical observations such as words and perceptions (Forrestal 2016, 565). Often the qualitative study does not start with defining a hypothesis to be proved or disproved in the study, or by defining specific variables to be examined (Creswell 2011, 16). Instead, the research questions are broad and open-ended, such as, "What motivated you to study health informatics?" Qualitative studies aim to analyze the data for emerging themes that can be used to derive meaning (Creswell 2011, 16). The nonnumerical data analyzed in these studies can be words, gestures, actions, pictures, audio and video recordings, activities, or perceptions of the study participants gathered from open-ended questions asking their opinions (Layman and Watzlaf 2017).

Usually, data in qualitative research are collected on a small group of participants via surveys, interviews, and focus groups (Creswell 2011, 16). A common way to analyze the

data is to look for themes in the respondents' answers to the research questions. In one qualitative study, the researchers collected data on the opinions on HIEs, e-prescribing (e-Rx), and personal health records (PHR) of individuals who were invited to participate in focus groups (Cochran et al. 2015). Qualitative analysis was conducted using transcribed audio data. For example, NVivo qualitative statistical software was used to identify common themes, which were compiled for each of the discussion categories: HIE, e-Rx, and PHR. Results indicated respondents were concerned about privacy and security, cost of implementation, and quality of care in these practices (Cochran et al. 2015, 1).

Mixed Methods Research

Mixed methods research studies use data collection and data analysis methods from both quantitative and qualitative research in one study. The major part of the mixed methods study can be either quantitative or qualitative. These studies are more complicated to undertake than a quantitative or qualitative study, but the combination of the two approaches provides a more in-depth picture of the phenomenon under study (Creswell 2013, 534). A common form of mixed method studies begins with a largely quantitative component followed by a smaller qualitative component such as a study to record customer satisfaction with 25 closed questions asking the user to rate their satisfaction on various factors followed by four or five written response questions for further comments and suggestions.

For example, a mixed methods study on the benefits of telemedicine in rural communities explored hospital staff perceptions on the benefits of telemedicine (Potter et al. 2016). Data were collected with surveys, phone interviews, and site visits. Descriptive and thematic data analysis was conducted. Survey respondents reported tele-emergency was beneficial to rural communities, and interviewees indicated that telemedicine improved the hospitals' image in the community (Potter et al. 2016, 1).

Experimental Research

Random sampling is a sampling method used to select an unbiased group of study participants (AHIMA 2017). **Experimental research** is conducted using random sampling to assign participants to either a study or a control group (Forrestal 2016, 585). During the study, one group is subjected to the study treatment, and the other group, called the **control group**, is not given the study treatment. It should be noted that administering treatment, or an intervention, can mean providing a new type of physical therapy, having participants listen to educational videos on coding practices, or giving them a medication. By contrast, the control group in these studies would use the existing type of physical therapy, would not listen to the video on new coding practices, and would be given a **placebo** that does not contain the new medication. Results from both groups are then compared to determine the effectiveness of the treatment. A **double-blind study** is one in which neither the participants nor the researchers know who is in the control group and who is in the study group. This type of study can minimize the influence from researchers, but it is not always feasible to conduct double-blind study, especially if the interventions for study and control groups are different. Experimental research can be used to give strong evidence of **cause and effect**, which means that researchers can state with some confidence that the variable under study caused, or did not cause, the effect under study (Layman and Watzlaf 2009, 115).

There are many **nonexperimental research** studies conducted where there is no random selection of participants, nor is there any random assignment of participants to study groups. Quasi-experimental, historical, descriptive, correlational, and qualitative are all nonexperimental studies. **Historical research** studies inspect past events like the events leading up to AHIMA's addition of the Certified Health Data Analyst (CHDA)

credential. Researchers conducting **descriptive research** collect and analyze data to describe the population characteristics (Huck 2012). In some big data studies, the entire population is included in the study. Therefore, there is no sampling, either. Data results typically include summary statistics (that is, mean, median, mode, range, and standard deviation) and frequency calculations for data characteristics (for example, percentage of males and females).

Correlational Research

The results of a **correlational research** study indicate the existence, direction, and strength of the relationships between two numerical variables (Huck 2012). Results of these studies can show that two variables are strongly or weakly related, or that they both increase together (positive relationship) or one increases but the other decreases (negative relationship). A **variable** is something that might have different values for different people or in different situations (Huck 2012). Correlational studies cannot be used to establish causation; only experimental studies can give strong evidence of cause and effect (Layman and Watzlaf 2017). With correlational studies, it could be that some variable outside the study caused the event. For example, researchers in a 2010 correlational study investigated the relationship between the patient's comprehension of their reasons for being admitted to the hospital from the emergency department (ED) and their overall satisfaction with the care received (Downey and Zun 2010). Results indicated that ED wait times were not a major determinant of their satisfaction with their care (there was a weak correlation between ED wait time and their satisfaction with their care), and that most would return to the ED if necessary.

Quasi-Experimental

Quasi-experimental studies are appropriate when the study variables cannot or should not be manipulated (Forrestal 2016, 586). This could occur because study variables cannot be changed (for example, age, race, gender) or should not be changed as they are integral to the study (such as patients with certain diseases, facilities that are using an HIE). **Retrospective research**, also called case control research, is quasi-experimental because the phenomenon under study occurred in the past. These types of studies are common in healthcare. Some examples of retrospective studies are researchers' examining postoperative infection rates at a facility, calculating the rate of coding errors for each medical coder, or evaluating the usage levels of EHRs from usage data collected over the last five years. Other types of quasi-experimental studies are field studies and natural experiments. A **field study** is done outside the laboratory or place of work. For example, researchers could have observed nurses' handwashing techniques in the field at a one-day surgery center. During a **natural experiment**, researchers have no control over the phenomenon under study. They do have the study planned but are waiting for the event to occur naturally, such as a study to evaluate the effects of an increase in physicians' co-payments on primary care physician visits (Jakobsson and Svensson 2016). The natural experiment research plan includes gathering existing data on the primary care physician visits then gathering data on the visits during the time frame of the study (when the higher co-pays are in effect) and comparing them. It is easy to discern that studies conducted to quantify the consequence of changing policies and procedures are also natural experiments, as are studies that measure the facility's process for restoring computer systems after a hardware system failure.

Clinical Outcomes Assessment

Healthcare outcomes are results, or what follows medical interventions by the providers of care. **Clinical outcomes assessment** is "research that seeks to improve the delivery of patient care by studying the end result of healthcare services," also referred to as

outcomes analysis (Forrestal 2016, 577). Many healthcare studies are conducted to assess clinical outcomes. The information gained by these studies is very important to improving quality of care along with controlling costs. The EHR affords new opportunities for tracking and assessing patient data. The Health Resources and Services Administration (HRSA), a resource for providers to aid in the development of their EHR, indicates that there are four steps for measuring clinical processes and outcomes.

- Identify the intervention or processes for improvement,
- Select the appropriate measures to evaluate the focus of the assessment,
- Establish a baseline of current practices using the selected measures, and
- Monitor and reassess the outcomes once the changes have been made to determine the effect on the outcomes (HRSA 2016).

The primary purpose of clinical outcomes assessment is improved patient care, which is accomplished by providing information for clinical decision-making and evaluating and improving the performance of clinicians. Clinical outcomes assessment data can also be utilized as justifications for pay for performance and to assist potential patients in the selection of a clinician.

IRB Processes and Procedures

Prior to conducting any type of research involving human subjects, researchers must obtain **institutional review board (IRB)** approval from the IRB office where the study is being conducted (HHS 2016). The IRB approval process provides review, oversight, and guidance to the researchers and is mandatory for federally funded programs (AHIMA 2017). The Office for Human Research Protections (OHRP) of the Department of Health and Human Services monitors compliance of IRB offices with the federal regulations governing research with human subjects. Most institutes have their own IRB compliance office and researchers might need to submit two IRB applications if they are not conducting the study at their place of employment. For example, an educator wishing to conduct research at a university hospital needs to apply to their university's IRB office and the hospital's IRB office. When applying for IRB approval, the researcher selects the type of IRB application that is exempted, expedited, or requires a full board review. Exempted applications would include using deidentified data such as past patient data with all identifiers removed (that is, deidentified computer report). Expedited submissions are used when the study potentially poses some risk to human subjects, and full board review and approval submissions deal with studies not in the first two categories or dealing with vulnerable subjects (21 CFR Parts 50 and 56: US Food and Drug Administration). **Vulnerable subjects** in the context of research studies are children, pregnant women, human fetuses, neonates, mentally disabled individuals, those who are educationally or economically disadvantaged, prisoners, or persons with incurable or fatal diseases (21 CFR Parts 50 and 56: US Food and Drug Administration). The study participants in these categories require additional consideration or protection. Whether the research is exempted or not, researchers must obtain informed consent from the subjects signifying that they understand the nature of the study, the study risks, and that the study data will be deidentified. Participants should be informed that they may opt out of the study at any time.

Check Your Understanding 7.1

Match the terms with the appropriate descriptions.

a. Research
b. Applied research
c. Quantitative research
d. Qualitative research
e. Experimental research

1. _____ Process aimed at discovering new information about a subject
2. _____ Tests a theory by defining measurable variables and collecting and analyzing data on those variables
3. _____ Practical or clinical research
4. _____ Research aimed at investigating nonnumerical observations such as images
5. _____ Research conducted by randomly assigning participants to an experimental or a control group

Answer the following questions.

6. Which type of research gives the strongest evidence of cause and effect?
   a. Experimental
   b. Nonexperimental
   c. Correlational
   d. Descriptive
   e. Predictive

7. Which type of research examines the existence, strength, and direction of the relationship between two numerical variables?
   a. Historical
   b. Evaluation
   c. Correlational
   d. Descriptive

Steps in the Research Process

There are fundamental steps that should be followed when embarking on research. For clarification, these steps are illustrated for a hypothetical quantitative correlational study. Before examining the research process, basic terminology related to the process steps must be reviewed. As mentioned previously, a variable is an attribute that may have different values. Quantitative studies have independent and dependent variables. As an illustration, consider the research question "Does providing additional training on a new EHR system increase the physicians' utilization of the system at this facility?" The **independent variable** is the variable that researchers believe will cause change in the **dependent variable**, which is the variable being measured. In the case of the EHR training, receiving the additional training (independent variable) is expected to cause a change in the level of utilization of the new EHR system (dependent variable), which is being measured. If the study results were graphed, the values on the x-axis would be the independent variables and the y-axis values would be the dependent variable, usually frequency values, because the value of y (level of utilization) is determined by the value of x (did or did not receive the training).

The research process steps are illustrated for a hypothetical quantitative correlational study. Initially, the quantitative researcher will generate a quantitative purpose statement, which is a sentence that describes what the study is examining, and identifies the independent and dependent study variables, the participants, and the study site. Research textbooks should be consulted when defining the purpose statement, as many provide

templates as guidelines (Creswell 2011, 123–124). For example, the following is a template for a purpose statement:

> The purpose of this study is to test (the theory) by relating (the independent variable) to (the dependent variable) for (participants) at the (research site). (Creswell 2011, 123)

For the correlational study, using the template, the purpose is to relate age (independent variable) with number of ED visits (dependent variable) for patients (participants) at Valley General Hospital (research study site; Creswell 2011).

Determine Research Question

Next, a quantitative **research question** will be generated. All quantitative studies have one or more research questions. The research questions restate the research problem and name the independent and dependent variables, participants, and research site, but with a narrowed focus. This study uses a relationship research question. The research question for this study is, "How does the age relate to the number of ED visits for patients at Valley General Hospital during the study period?"

page 174

Some researchers might also advance **hypotheses**, which are their predictions of the results that are often based on findings from their extensive literature review. The **null hypothesis** indicates there is no association between the independent and dependent variables (AHIMA 2017). The **alternate hypothesis** indicates there is an association between the independent and dependent variables (AHIMA 2017). For this study, the research hypothesis will be "The greater the age of the patient, the more ED visits they will have at Valley General during the study period." Formally, the null hypothesis is tested as "There is no difference between age and number of ED visits at Valley General in the study period." The alternate hypothesis is tested as "There is a difference between age and number of ED visits at Valley General in the study period."

Define Study Variables and Their Data Types

Structured data collected for a study can have four levels of measurement: nominal, ordinal, interval, or ratio (AHIMA 2017). Nominal and ordinal data are expressed in categories, whereas ratio and interval data are continuous data (White 2016, 512). The selection of data measurement levels for data representation is an important part of a study because it determines what type of statistical tests can be conducted.

For this study, the independent study variable is age and the dependent variable is number of ED visits. To further define age, the researcher must decide how to represent the data for age. The researcher could use categorical ordinal age groups like 20–25 years, 26–30 years, and so on. Alternatively, age could be recorded as a continuous number such as 20.5 years or 48 years. To simplify data collection, age will be a continuous variable and number of ED visits will be a discrete counting number. Studies also typically collect demographic data about their participants, so the researcher will add two demographic variables to collect race and gender, and they will be nominal variables.

Identify Sample and Population

The **population** is the universe of phenomena, objects, people, or data under investigation from which a sample is taken (AHIMA 2017). A **sample** is a group of patients selected to represent the population and that sample is tested using **inferential statistics** techniques that allow the researchers to make statements about the population's characteristics based on the sample's characteristics (AHIMA 2017). For this study, suppose the ED manager says that her staff is very busy but she will try to provide the sample data for the study by having the staff collect data on 30 patients a day, with each shift surveying 10 patients. Because it is always possible that the study sample is not

representative of the population, at this point data should be obtained on the age (or age groups), gender, and race of all the ED patients seen during the study period.

Data Collection Methods

Types of data collection methods include census, sample, questionnaires (survey), observation, and interviews. These are defined as follows.

- Census—A survey that collects data from all members of a population
- Sample—A subset of the population selected for study
- Survey—Self-reported research where the individual is the data source
- Observation—Method of collecting data by watching research participants
- Interview—Member of the population being studied responds to oral questions (AHIMA 2017)

In the example, the study period is determined to be one year. Survey data can be collected by emailing the survey to participants, calling the participants on the phone, interviewing the participants, or sending them a link for a web-based survey. For this study, a paper survey is used to collect data in the ED, and retrospective record reviews are used to determine how many visits the patients in the sample have during the study year. Researchers also need to decide how to handle return visits by people who have already been surveyed. Researchers should decide if they want to collect a new form for previously surveyed participants returning to the ED. For accuracy, data could be collected each time a participant comes back. However, unless this is communicated to the data collection staff in the ED, valuable data could be lost.

Data Collection Instruments

The data collection instruments provide the means for consistency in obtaining the data in a required format (AHIMA 2017). For example, an instrument could be a web-based online survey, or a paper form for the respondents to fill out. This is important because data collection issues, such as failing to define the study variables, can result in unusable survey responses. For this study, a query will be run to export the data for age, ethnicity, gender, date, and time of the ED visit from the electronic ED registration system. The data will be collected by the ED staff for at least 10 patients on each shift. Researchers will enter the deidentified data into a Microsoft Excel spreadsheet. In this way, the online survey can facilitate obtaining structured survey data for analysis.

Research Electronic Data Capture (**REDCap**) is a web-based platform that supports different types of clinical and translational research, such as clinical trials, retrospective studies, and building and managing surveys and databases (Harvey 2018). REDCap was developed by a small group of clinical researchers at Vanderbilt University in 2004 to facilitate single and multisite research studies. Due to its lightweight and easy-to-use interface, REDCap has grown substantially to become a widely used research data collection and management software solution. Researchers can obtain free access to the REDCap software at no charge from nonprofit organizations that are REDCap Consortium Partners. If the institution has already joined the REDCap Consortium, the researcher will need to contact the institutional REDCap administrator to request a user account. Many health institutions also provide seamless connection between the REDCap software and the institutional EHR system through the Fast Healthcare Interoperability Resources (FHIR) web service so that the researcher can pull structured EHR data directly into the REDCap database. Since REDCap leveraged data standards, this service has greatly improved interoperability with health records and fostered clinical and translational research.

Check Your Understanding 7.2

Mark the answer as true or false.

1. The independent variable is the variable being measured.
2. The dependent variable is the one the researcher believes will cause a change.
3. Hypotheses are predictions of the study outcome by the researcher.
4. The sample is a subset of the population of interest in the study.
5. Survey data may be collected by emailing the survey, calling participants on the phone, interviewing participants, or sending them a link for a web-based survey.

Answer the following questions.

6. Which is the data collection method for self-reported research where the individual is the data source?
   a. Survey
   b. Observation
   c. Interviews
   d. Census
7. What is not a characteristic of the alternate hypothesis?
   a. It is a statement of the researcher's working hypothesis
   b. It is a prediction of the study outcome
   c. It indicates no association exists between the independent and dependent variables
   d. It indicates an association exists between the independent and dependent variables

Data Analysis

Data analysis is defined as "the methods that help to describe the facts, detect patterns, develop explanations and test hypotheses" (AHIMA 2017). To conduct data analysis, the reliability and validity of the data must be evaluated. The statistical analysis software must be selected or purchased, and the statistics tests must be selected. After these preliminaries, descriptive and inferential statistics can be generated using the statistical software and the correct statistical tests.

Reliability and Validity

Research should be valid and reliable. **Reliability** of instruments indicates they give similar results on repeated administrations of the test. **Validity** is the extent to which an instrument measures what it is supposed to measure (AHIMA 2017). **Internal validity** is a characteristic of the study design (Forrestal 2016, 588). Threats to internal validity come from factors inside the study that interfere with proving that the study outcome is the result of the independent variable (Forrestal 2016, 588). Common threats to internal validity are due to

- History—A current event occurring during research affects the outcome
- Maturation—Subjects mature during study
- Testing—Taking repeated tests affects scores
- Instrumentation—Lack of consistency in data collection or changing instruments
- Statistical regression—Subjects selected due to extreme scores tend to have scores that regress toward the mean with repeated observations

- Differential selection—Individuals selected with certain characteristics for control and experimental groups
- Attrition of subjects—Subjects leave during the study
- Diffusion of treatment—Participants learn who is in the control group and who is in the treatment group (Forrestal 2016, 588).

External validity threats come from outside the study, such as the selection bias where the characteristics of the participants are so narrow that results cannot be generalized, or the setting characteristics of participants that prevent generalization. For example, a study of the challenges of underprivileged teenagers growing up on a Caribbean island would have selection bias due to the specific selection characteristics and the unique setting that would prevent generalization to all teenagers in the United States.

If the study instrument is the Excel data collection worksheet then validity could be established by having the instrument reviewed by an experienced researcher, and by conducting a pilot study using the instrument. To establish internal validity, the researcher must control for confounding variables that could influence the study outcome, which could be done by controlling for the common threats to internal validity, such as making sure the data collection procedures are consistent and not selecting subjects with extreme scores. With respect to external validity, the sample will be compared to the population of ED visitors to determine if it is representative. However, it is still a small sample from one ED for only one year, so this will be listed as a study limitation.

Select Statistical Analysis Software

Commonly used statistical software includes Microsoft Excel, Statistical Package for the Social Sciences (SPSS), SAS, R software, Stata, and Python. The choice of software used by the researcher may be limited by the research budget and the lack of availability of more expensive software packages. Specifically, R and Python are free and widely used in health informatics, Excel is moderately priced, and the remaining statistical analysis software solutions require expensive licensing that must be renewed each year. Nevertheless, for large data sets the licensed software provides many advantages in terms of usability, ease of running the tests, and results output in a more usable format, such as tables and diagrams.

page 178

Microsoft Excel

Microsoft Excel is a spreadsheet product that comes with the installations of Microsoft Office for Windows or Macintosh computers (Microsoft 2023). There is a free Excel add-in that can be enabled to easily conduct statistical tests. Many web-based survey tools, such as SurveyMonkey and Qualtrics, offer users the ability to export their data to Excel. Excel is a good choice for analyzing smaller data sets. See figure 7.1 for an example of an Excel datasheet.

SPSS

The Statistical Package for the Social Sciences (SPSS) is an IBM product consisting of a family of products for data analysis, reporting, import, and export. SPSS is widely used in research and it supports imports and exports from a large variety of systems in many different formats (SPSS 2023). There are three editions of SPSS: Standard, Professional, and Premium. The Standard edition provides the basic analytics that would be used for most projects. The Premium version is designed for enterprise-wide data analysis needs. SPSS usage requires purchase of a license each year. SPSS can be used to perform sophisticated statistical tests such as factor analysis; binomial, multinomial, and logistic regression; and predictive analytics. SPSS has customers in the medical, academic, insurance, banking, government, manufacturing, and survey and marketing fields.

page 179

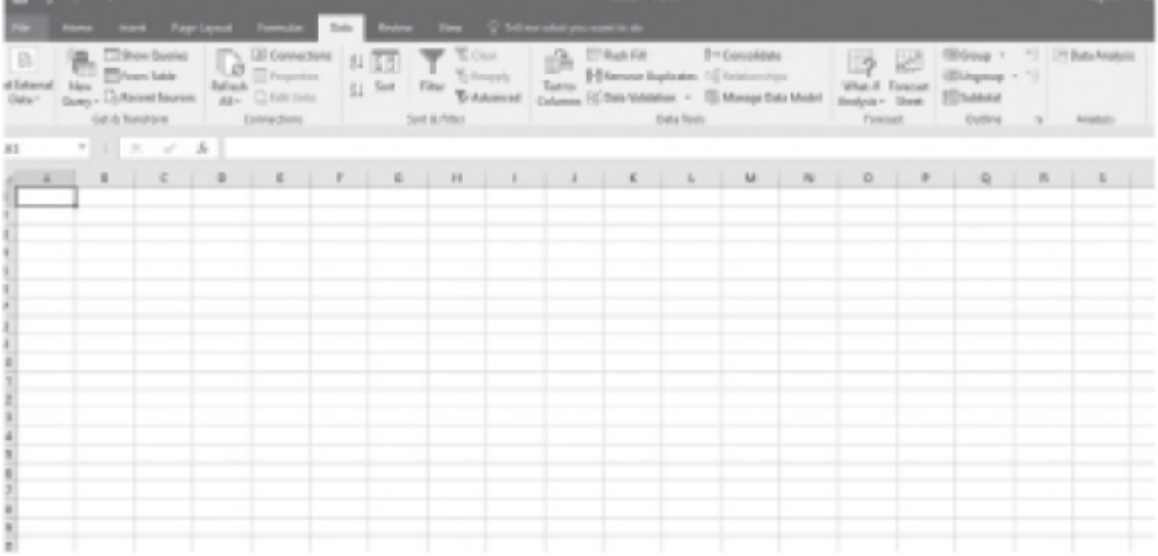


Figure 7.1. Excel datasheet

▶ Expand/collapse Extended Description

Used with permission from Microsoft.

SAS

SAS®, Statistical Analysis Software, has a large set of integrated enterprise solutions for performing advanced analytics, business and customer intelligence, data, decision, performance, and risk management. SAS has a large programming language set that can be used to load and analyze a large variety of data (SAS Institute Inc. 2023). It is used globally and has a large user group. Uses in healthcare include studies on readmission, health outcomes, and patient safety.

R software

R statistical software is a free open-source software program, based on the R programming language developed at Bell Labs, which can import and export data to a large variety of statistical software (R Development Core Team 2023). Currently, there are many free resources for learning R (CRAN 2023). The software can be downloaded and installed on the user's computer with versions available for Windows, Mac OS, and Linux operating systems. R has a significant learning curve because many of the commands are entered at a command prompt, which might be unfamiliar to users. However, there are many tools to assist the user such as a development environment called RStudio that can be downloaded (RStudio 2023). There is also an active R user community (Jackson 2014). In healthcare, R has been used to explore duplicate record sets (Haenke Just et al. 2016), and to compare EHR incentive program attestations from rural and urban facilities (Sandefer et al. 2015). Figure 7.2 shows the RStudio interface.

Stata

Stata data analysis and statistical software was developed by StataCorp, which also publishes a peer-reviewed quarterly statistical journal focused on Stata use (StataCorp 2016). Stata provides basic and advanced statistical analysis and has a rich set of data graphing tools. There are Stata products available to handle small, moderate, and large data sets on Windows, Mac, and Unix platforms. Uses of Stata in healthcare include predicting 30- to 120-day readmissions for Medicare patients (Ostrovsky et al. 2016) and

facilitating the development of innovative predictive models developed for cardiac prevention studies (Barkhordari et al. 2016).

Python

Python is a high-level, general-purpose, object-oriented programming language that has been widely used by data scientists. Due to its easy-to-learn features and rich resources of libraries and tools, Python has become one of the most popular data science programming languages. There are a number of Python libraries, three of which have been used extensively for data science tasks. NumPy library is commonly used for processing large dimensional data arrays; Pandas library provides convenient data manipulation and analysis functions; and Matplotlib is a powerful tool for data visualization. In addition to these data analysis features, Python is particularly good for development and deploying machine learning and artificial intelligence models and algorithms. Two major deep learning tools, TensorFlow (2022) and PyTorch (Linux Foundation 2022), are also developed in Python. For example, researchers (Attia 2019) used TensorFlow to develop an artificial intelligence (AI) model that used electrocardiograph (ECG) to detect the electrocardiographic signature of atrial fibrillation present during normal sinus rhythm. More and more healthcare organizations have established institutional research computing environments that support data analysis in Python.

Figure 7.2. RStudio interface

▶ Expand/collapse Extended Description

Determine Statistical Formulas to Use or Statistical Tests to Conduct

To select the correct statistical test for inferential statistics, the independent and dependent variables' data types and numbers must be considered, as well the purpose of the study. A statistical textbook should be consulted, as failing to select the correct test

can affect study validity (White 2016, 513). For example, in the current example of age and number of ED visits there are two variables (bivariate), both of which are numbers. The aim is to determine the existence and strength of a relationship between the two variables, so Pearson's correlation is an appropriate test. Pearson's correlation generates a value between −1 and +1 that indicates the strength and direction of the correlational relationship. For example, a value of 0 indicates no relationship, whereas values of −1 and +1 indicate strong negative and positive relationships, respectively.

Compute Descriptive Statistics

Descriptive statistics are used to analyze study variables like age and number of ED visits because they can describe the features of the data, but they are not used for hypothesis testing. In many studies the descriptive (summary) statistics generated are frequency distributions, and the mean, median, mode, standard deviation, and variance.

Measures of Central Tendency

Specifically, continuous variables can be analyzed using measures of central tendency, which are the mean and median (White 2016, 520). The mean is the arithmetic average, which is computed by adding up the values in the data set and dividing by the number of values in the data set. The median is the middle value, which is computed by first arranging the values from smallest to largest, then either picking out the middle value (if the number of values is odd) or averaging the middle values (if the number of values is even).

Outliers are values that have very high or very low values in comparison to the rest of the values in the data set. Outliers can influence the mean to "move" toward the outliers. If an outlier is very high, the mean will be artificially high. In this case the researcher might wish to report the median instead of the mean because the median is not influenced by outliers (White 2016, 520).

The mode is reported for categorical data. It is the most frequently occurring value. A data set could have no mode, one mode, or many modal values. For example, the data set {1, 2, 3, 4, 5, 6, 7, 7, 7} has one mode, which is the value 7, because it occurs the most frequently (three times).

Measures of Spread

For continuous variables, the variance, standard deviation, and range are used to calculate the spread of the data, which is the distance from the mean of the data set. Variance is computed as the average of the square of the distance from the mean. The standard deviation is calculated by taking the square root of the variance. The range is the difference between the largest and the smallest value in the data set. For example, the range for the data set {1, 2, 3, 4, 5, 6, 7} is 7 − 1, or 6.

Compute Inferential Statistics

Inferential statistics can be computed using rates and proportions or simple linear regression or by computing a value for the strength of a relationship between two variables or with other advanced statistical tests (White 2016, 526–530). For the example under consideration in this chapter, a bivariate correlational test can be conducted in Excel or SPSS to determine the strength and relationship between age and number of visits to the ED per year.

Explicitly, the statistical test would be a Pearson's correlation test, which in Excel is calculated with the CORRELL function. This function generates an r correlation coefficient, which can be 1 (strong positive relationship), −1 (strong negative correlation), or somewhere in between. In determining the strength of the relationships, the absolute value of r is utilized and the closer to zero the weaker the relationships. For example, a

Pearson's correlation r value of −0.98 indicates a strong negative correlation, a value of 0.85 indicates a strong positive correlation, and a value close to 0 indicates no correlation. The conditions for the test are met because both age and number of ED visits are numbers, and data have been collected for both variables.

Similarly, in SPSS the data must be prepared, which means that initially variables loaded into SPSS need to be coded and labeled. Once the coding is done, SPSS has many built-in functions and features that automatically generate output and graphs (Green and Salkind 2013, 69). After loading the data, the user runs a bivariate correlation on age and number of ED visits. The SPSS output window will display a correlations table containing the Pearson's r correlation value, which is .672 and the significance will be given by Sig. (2-tailed) as .098. A significance value greater than .05 indicates there is no statistically significant correlation between the variables. A significance value less than or equal to .05 indicates there is a statistically significant correlation between the variables, and variables that are related will increase or decrease together. For this test, the significance is .098, which is much greater than .05 so the variables age and number of ED visits are not significantly correlated. The conclusion is that age is not related to the number of ED visits.

Check Your Understanding 7.3

Answer the following questions.

1. Which of the following description match the external validity?
   a. Threats to the study like poor data collection instruments
   b. Test instrument gives consistent results on repeated administration
   c. Threats to the study such as selection bias
   d. Extent to which an instrument measures what it is supposed to measure

Indicate whether each of the following is true or false.

2. SAS is a Microsoft product that can be used for data analysis and statistics.
3. SPSS is an IBM product widely used in research.
4. Python is a free open-source software program.
5. The type of statistical test is dependent on the data types of the independent and dependent variables, the number of independent and dependent variables, and the purpose of the study.

Advanced Data Analysis

Advanced data analysis is growing in popularity due to the amount of big data now being analyzed in healthcare. Advanced data analysis includes trend, forecast, and predictive analysis, as well as data mining and epidemiological applications of data analysis. Big data are data sets so large that new tools for analysis are necessary to explore them (Patena 2016, 750).

Trend Analysis

Trend analysis can be conducted with scatterplots (Green and Salkind 2013, 236), t-tests (Green and Salkind 2013, 145), or one-way analysis of variance (ANOVA) testing (Green and Salkind 2013, 163). **Scatterplots** are graphs that show correlation between variables; the x-axis displays the independent variable and the y-axis shows the dependent variable. Scatterplots are generated by most correlation tests, and the values of 0, −1, and +1 indicate no relationship, perfect negative relationship, and perfect positive relationship,

respectively. Scatterplots are excellent visualization tools and are present in many research studies. Figure 7.3 shows a scatterplot for hypothetical data on age and length of stay in the hospital. Visual examination of this scatterplot shows an upward trend indicating that as age in years increases so does length of stay in the hospital.

A **t-test** compares a mean to a test statistic to determine how they are related. It can be used to compare current data to a company benchmark. For example, is the current average ED wait time more or less than the facility's desired value of 40 minutes for ED wait time? Trends can also be determined using an **ANOVA** test to compare two or more means to see if they are equal, or if at least one of them is different. There are several types of ANOVA tests. A commonly used ANOVA test is the one-way between-subjects ANOVA, which should be used when there is one independent variable and one numerical dependent variable (Huck 2012, 217). It is called one-way because only the levels of one factor, the independent variable, are used when putting the participants into the three groups. For example, if a participant takes a high, low, or no dose of Vitamin C then the three levels of this factor are used to create the three groups according to level of Vitamin C usage. The ANOVA test could then be conducted to determine if there was any difference between the groups with respect to number of days to develop cold symptoms (the dependent variable; Green and Salkind 2013, 163).

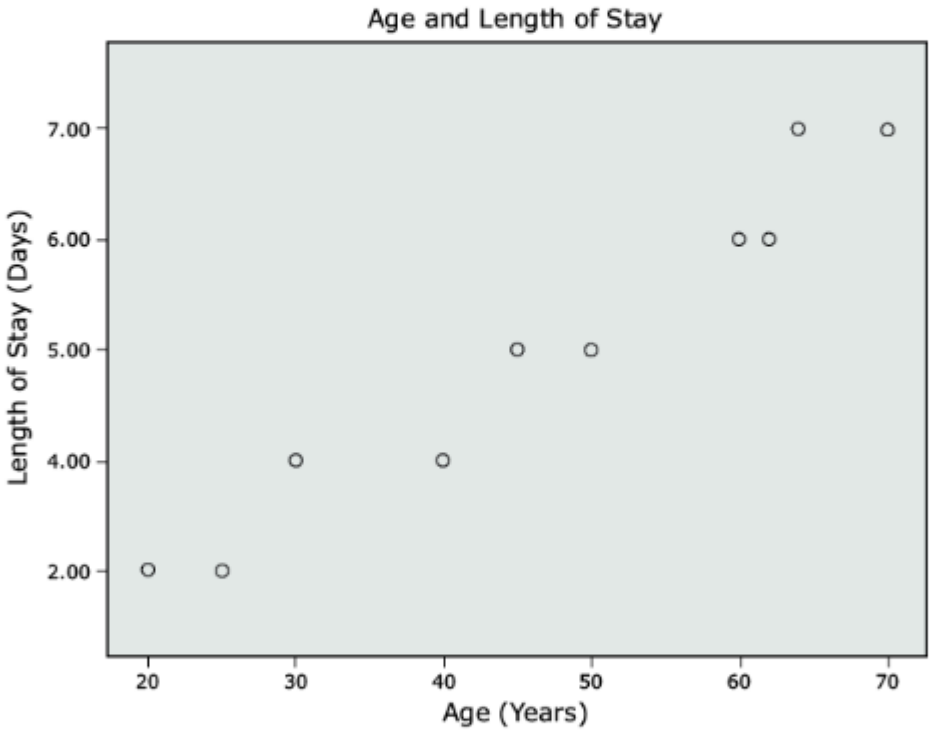


Figure 7.3. Scatterplot for age and length of stay

▶ Expand/collapse Extended Description

The ANOVA test can indicate that two or more of the Vitamin C groups differ significantly, but it does not indicate which of the two groups are different (Green and Salkind 2013, 167). To determine which groups are different, an additional post hoc test must be conducted. For example, Dunnett's C test is a post hoc comparison for the one-way ANOVA test that is available in SPSS (Green and Salkind 2013, 168).

Predictive Analysis

Predictive modeling is "a process used to identify patterns that can be used to predict the odds of a particular outcome based on observed data" (AHIMA 2017). These input data are usually historical as well as current data. **Predictive analysis** predicts future outcomes and is often used for simulating future scenarios that help managers optimize their outcomes. Researchers in a study designed to evaluate predictive and descriptive analysis for use in optimizing nursing staffing patterns at a small rural hospital determined that their predictive tools helped managers to maintain optimal nursing staffing levels (Ramsey 2014). The use of predictive analytics is increasing because it can offer managers a way to model what-if scenarios to select the best paths (Minsker 2015).

Forecasting

Forecasting is calculating or predicting some future event or condition through study and analysis (AHIMA 2017). Forecasting in healthcare predicts health disease episodes or health events like number of admissions, life expectancies, or disease outcomes for one individual (Soyiri and Reidpath 2013). Forecasting is a commonly conducted test. For example, SPSS has a Forecasting selection on the drop-down menu. Forecasting can be performed using several mathematical models such as time series analysis, where data collected over time are analyzed to find trends, cycles, and seasonal changes in the data.

One study described how forecasting was utilized to determine peak asthma times (Soyiri et al. 2013). First, two years of retrospective data were collected from the National Health Service in England on hospital admissions where the primary diagnosis was asthma and the admissions occurred between January 1, 2005, and December 31, 2006 (Soyiri et al. 2013, 2). The binary dependent variable was days of peak demand for admissions for asthma; that is, days when there were 40 or more asthma admissions (Soyiri et al. 2013, 2). Data were analyzed with multiple regression to develop a multivariate model to predict the days of peak demand for asthma visits to the ED.

Data Mining

Data mining is "the process of extracting and analyzing large volumes of data from a database in order to find patterns in the data" (AHIMA 2017). Tools for data mining include IBM SPSS data modeler, Oracle data mining, SAS Enterprise Miner, HP Vertica Analytics Platform, and Microsoft Analysis Services. For smaller data sets, the R statistical tool has data mining packages. Data mining has been used to explore family history (Hoyt et al. 2016), to explore factors associated with decubitus ulcers (Dheeraj et al. 2015), and to identify fall-related injuries (Luther et al. 2015).

Artificial neural networks (ANNs) are "predictive computer models that use interactive learning to process multiple inputs in order to detect patterns" (AHIMA 2017). Thus, their behavior is similar to the neurons of the human nervous system. ANNs have been used in radiology, cardiology, and neurology (Hoyt et al. 2016, 2). Regarding data mining, researchers used ANNs to analyze digital medical family history data on approximately 2,400 participants (Hoyt et al. 2016, 2). Results from data analysis with ANNs were compared to previous cross-tabulation results for the same data, and they were determined to be sufficiently correlated, indicating that ANNs are useful tools for detecting patterns in data.

Researchers aiming to develop an improved model for predicting decubitus pressure ulcers used R statistical software to compare the results from four data mining models on four years of data (Dheeraj et al. 2015). The data mining models examined were logistic regression, decision trees, random forests, and multivariate regression splines. Results indicated that the random forest model, a special type of decision tree, was the most predictive of pressure ulcers for this data set (Dheeraj et al. 2015).

Statistical text mining (STM) was utilized to identify fall-related injuries in EHRs (Luther et al. 2015). Data were collected for visits related to falls (n = 1,653) at the Veterans Health Administration (VHA) ambulatory care clinics. Data from the VHA EHRs were extracted for two days after the ambulatory clinic fall visit, and these data were annotated by three clinicians to mark text related to the falls. A statistical software package called RapidMiner was used to develop a statistical model that correctly predicted 87 percent of the visits for falls (Luther et al. 2015). This result is important because using data mining, persons at risk for falls can be more easily identified and offered fall-prevention information (Luther et al. 2015).

Tools have been developed to support clinical text mining and clinical natural language processing (NLP). cTAKES is an Apache open-source project built on the Apache Unstructured Information Management Architecture (UIMA) framework for extraction of information from unstructured clinical notes (Savova et al. 2010). MetaMap, developed and maintained by the National Library of Medicine (NLM), aims to extract and map biomedical text to the Unified Medical Language System (UMLS) Metathesaurus or Concept Unique Identifiers (CUI) code (Aronson et al. 2008). MedTagger is developed by the Open Health Natural Language Processing consortium that provides an open-source clinical NLP pipeline based on the UIMA framework for rule-based information extraction from clinical text (Liu et al. 2013). The Clinical Language Annotation, Modeling, and Processing (CLAMP) tool is another clinical NLP toolkit that provides state-of-the-art clinical NLP components for information extraction as well as a user-friendly graphic interface (Soysal et al. 2018). MedSpacy is a newly developed Python library for performing clinical NLP and text processing with the spaCy framework (Eyre et al. 2021). It leverages the state-of-the-art functionalities and packages in Python and could be easily adopted by people who are familiar with Python.

Epidemiological Applications

Epidemiological studies are concerned with finding the cause and effect of diseases (AHIMA 2017). Public health informatics involves studying epidemiological data for large numbers of people around the world. Finding a data mining tool that works in remote locations can be challenging due to lack of reliable internet connectivity. R has an EpiTools package that provides basic programming for epidemiological problems. EPI-Info is freeware available from the Centers for Disease Control and Prevention (CDC) that facilitates data collection, analysis, visualization, and reporting (CDC 2016). Using this software, advanced statistical analysis and geographic information system mapping can be quickly and easily accomplished. Additionally, it works well in remote locations with sporadic internet connectivity (CDC 2016).

Check Your Understanding 7.4

Match the terms with the appropriate descriptions.

a. Scatterplots
b. Predictive analysis
c. t-test
d. ANOVA

1. _____ Graphs that show correlation
2. _____ Compare two means of two data sets to determine the relationship (c)
3. _____ Compare means of four data sets to determine the relationship
4. _____ Calculating the odds of an outcome

Answer the following questions.

5. Data mining is
    a. The process of extracting large volumes of data to look for patterns
    b. A test for bivariate relationships in quantitative variables
    c. An epidemiological application for analysis
    d. a and b
6. Health forecasting cannot predict
    a. The number of admissions in the future
    b. Life expectancies
    c. Disease outcomes
    d. The number of admissions last year
7. Which of the following is not a data mining tool?
    a. SPSS Data Modeler
    b. SAS Enterprise Miner
    c. Microsoft Analysis Services
    d. Microsoft Excel

Data Visualization

Data visualization refers to the representation of data pictorially with graphs and other visual displays. Representing the data with graphs helps the researcher to get an overview of the features of the data. Viewing graphs, it is easier to detect outliers, spot trends, and see the skew of the data. When selecting the type of graph, recall that the data type determines what graphs can be created. If categorical data were used, the bar chart, pie charts, and frequency tables are appropriate.

Microsoft Excel

Under the Insert tab, Microsoft Excel provides a large array of graphs. After selecting the data, the user can decide to use the Recommended Charts, or select from pie, bar, line or PivotCharts. Pie charts can be created to show the percentage of the whole, such as the percentages of medical specialties in the hospital (for example, OB/GYN, MED/SURG, Cardiology). A line chart is a good choice for displaying large amounts of data over time. PivotCharts can be used to summarize large amounts of data according to several factors. For example, charges can be summarized by age and length of stay. Figure 7.4 shows a PivotChart for total charges by payer class and age group.

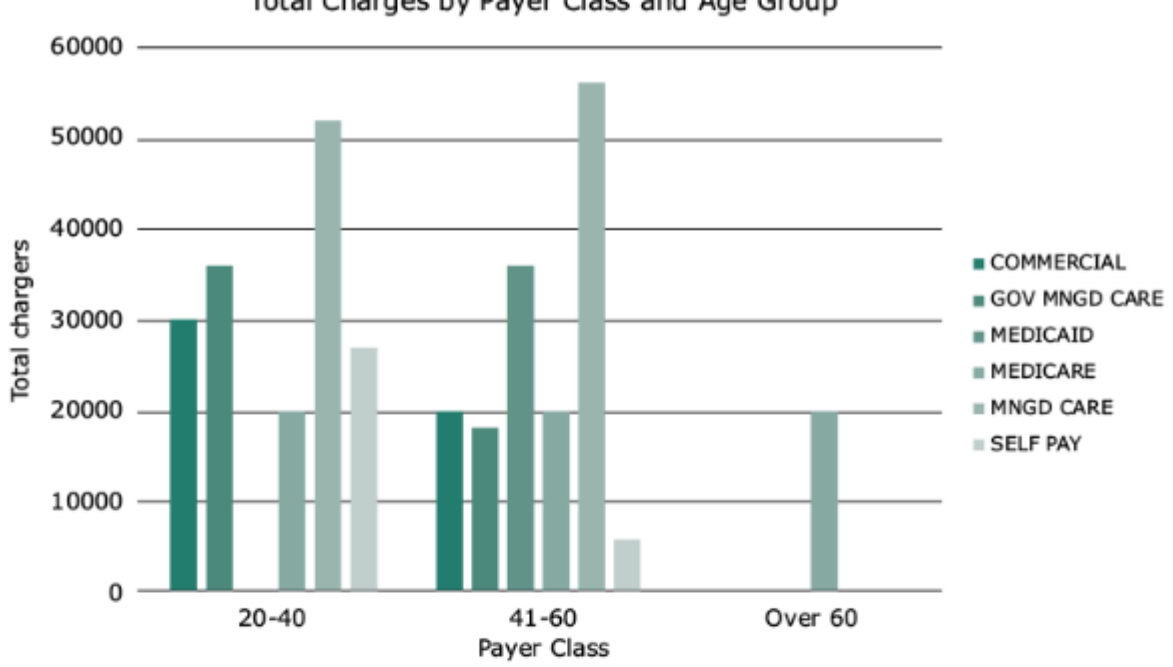


Figure 7.4. PivotChart of total charges

▶ Expand/collapse Extended Description

Used with permission from Microsoft.

SPSS

In addition to the aforementioned basic graphing suite, SPSS offers geospatial analytics, bulk imports for large data sets, and reports that can be posted to the web. SPSS offers visualization custom chart examples with details on how to perform the customization at their website (IBM Corp. 2023). The IBM SPSS Modeler has added functionality that allows users to create geospatial maps using their data. The visualization can be modified by changing the coordinate system, displaying the data in layers, or generating predictions of what will happen at a given location in the future such as disease outbreaks.

R Software

Using R, the user can often create graphs with one line of code. In R, the command plot() can be used to create a scatterplot. There is also a more advanced graphics package called ggplot2 that can be installed to generate univariate and multivariate graphs for numbers and categorical data in addition to the basic graphs available with the built-in package.

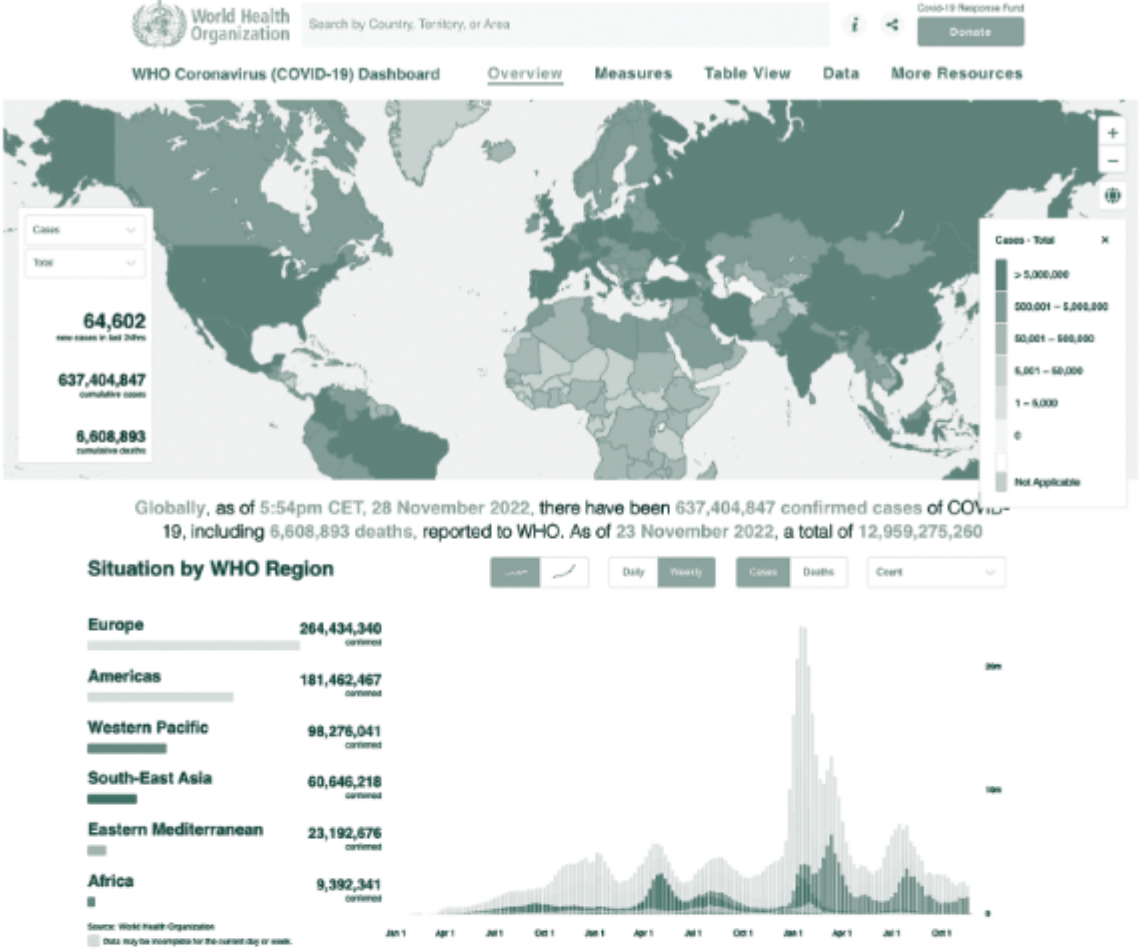


Figure 7.5. WHO COVID-19 dashboard as of November 28, 2022.

▶ Expand/collapse Extended Description

Used with permission from WHO.

Python

There are many visualization libraries available for Python, but Matplotlib has been the most widely used library for plotting. Many data analysis libraries, such as Pandas, have plotting functions based on Matplotlib. The latest release of Matplotlib is 3.5.1 as of 2022. Seaborn is another commonly used visualization library that produces more aesthetically modern plots than Matplotlib. Plotnine is a Python version of ggplot2 in R. The Gioplotlib toolbox is usually used for creating maps and plotting geographical data.

Dashboards

Dashboards are web-based portals that allow the user to display statistics and data from multiple sources on one screen, which allows the data relationship to be more easily visualized. Dashboards can be programmed to code areas that are failing (with different colors) so that the user can quickly identify problem areas. Dashboards are becoming more popular in healthcare and also are present on many government websites. Figure 7.5 presents World Health Organization (WHO) Coronavirus disease (COVID-19) from their website. Health data scientists sometimes also use Python libraries, such as Dash and Holoviz, to build a visualization dashboard to create a more interactive interface.

Check Your Understanding 7.5

Match the terms with the appropriate descriptions.

a. Excel
b. SPSS
c. R software

1. _____ Microsoft spreadsheet tool that produces PivotCharts
2. _____ Tool with Data Modeler for creating geospatial graphs
3. _____ Free software that can create graphs from one line of code Indicate whether each of the following is true or false.

Indicate whether each of the following is true or false.

4. Dashboards are portals that allow the user to display statistics and data from multiple sources on one screen.
5. Excel integrates easily with R software.
6. A line chart is a good choice for displaying percentages of the whole.

Data Retrieval

Data retrieval is the process of obtaining data from a healthcare database (AHIMA 2017). Initially, researchers collect data that are then analyzed and stored in a data repository. This repository could be a part of one of the statistical software packages mentioned thus far, or it could be in a database system like Microsoft Access, Oracle, DB2, or SQL Server. After the data are stored, users want to get the data out of the system and use them for reporting, or export them to another system. The process of getting the data out of the system is called data retrieval. To accomplish data retrieval, data languages and tools must be used.

page 190

Languages and Tools

Data can be imported and exported from statistical software and other data repositories using built-in tools. Data are loaded into the system with the built-in **import** function. This can be done in one step by selecting the import function then picking the correct data type from the ones listed on the statistical software's drop-down menu. Common data types for import are text or comma delimited files. To get the data out of one system, the **export** functionality is used. For example, in SPSS, the data can be exported in a variety of formats such as an Excel file or a Microsoft Access database. Other databases could be connected as well, or the user can save the data file and import it manually to the other system. For example, the SPSS Save Data pop-up menu lists a variety of file types that can be imported such as SPSS, tab or comma delimited, Excel, dBase, Stata, and SAS.

SQL: Microsoft Access, Oracle, IBM DB2, SQL Server

The most common language used for data retrieval is **Structured Query Language (SQL)**, which is used to query relational databases to retrieve the data for display. The basic principles for using SQL commands with databases were developed by IBM (Dolezel 2015a). The syntax for the data retrieval commands is available in the database documentation, and there are numerous resources that can assist the beginning database reporter. Of course, the syntax of these statements used for retrieval will differ across the systems, but Access, Oracle, DB2, and SQL Server all use a query language, which is a language used for retrieving data from a database, and the retrieval statements are similar. One SQL data retrieval query can be used to gather data from one or more tables located

in one or more databases, and these databases could be in another building, town, or country. They can be small or large, or even, with some bridging tools installed, from a different database vendor. Please refer to Chapter 5 for more details regarding SQL.

SPSS Data Explore Function

SPSS has a useful procedure called the Explore function. This procedure generates summary statistics and boxplots, stem and leaf plots, and normality and spread plots for quantitative variables. The advantage is that a variety of statistics are available using one procedure. This is useful for experienced users as they can quickly check for data skewness, examine normality tests, and glance at plots to see the shape of the data. Because SPSS sends all output to the output window, the results of running the Explore procedure can easily be saved for later review. However, the beginning SPSS user might find the multipage output to be overwhelming, so more value could be derived from doing these tests individually.

Check Your Understanding 7.6

Indicate whether each of the following is true or false.

1. Popular database systems, such as Microsoft Access, Oracle, and DB2 use SQL to manipulate database.
2. SPSS has an Explore function that will generate summary statistics and several useful graphs.

Data Reporting

Healthcare data analysis typically generates multiple reports that can be evaluated for performance monitoring, trend analysis, and other related quality improvement and management activities. Microsoft Excel is often used for graphing and to create Pivot tables. Users with large data sets or databases in several geographic locations might wish to purchase a commercial reporting system.

Tables, Charts, and Figures from Statistical Software Results

Table, charts, and figures from statistical software results can be exported to PDF files, Microsoft Word documents, or Microsoft Excel. The exported data can be used for reporting or as input for further analysis in another system. For example, Excel data on length of stay by diagnosis-related group can be imported into SPSS and analyzed to determine if any trends exist. Users wishing to select just one graph for use as a document or Microsoft PowerPoint will find this ability present in almost all of the statistical software packages. In the SPSS output window, chart images are selectable and they can be copied as images and pasted into a Word document. Excel graphs, tables, and PivotCharts are easily copied to a document as well.

Reporting Software

Many database management systems have their own built-in reporting software. However, most of these systems are designed to be used by the database administrators for monitoring and maintenance. For that reason, an important first step before writing any code or using any reporting tools is understanding what data and databases a facility has available. This can be found in the metadata catalogs for the database system.

Metadata is information about the data in the database, typically maintained at the system level and the user level (Dolezel 2015b). It is essential that the user contact the database administrator (DBA) who is assigned to their project prior to designing a report and work with them to explore the metadata. Metadata information includes the names of all tables, objects, fields, indexes, primary and foreign keys, usernames, and privileges for the database. To create a report, it is best to work with the DBA to identify the table names and fields needed to query for your report. The DBA can also help informatics specialists to optimize queries and in particular the queries where data comes from two or more tables, called table joins, giving attention to not causing the database to become bogged down.

page 192

Users with large databases and heavy reporting needs often decide to purchase a separate reporting system from a vendor such as Crystal Reports by SAP, Microsoft SQL Server Reporting Services, and IBM's Hyperion system. Database reporting can also be conducted using Microsoft Access reporting wizards and the SQL command line. The PC user can import Excel data for a report and print or save the report output.

Presentation Software

Presentation software is used to disseminate ideas at conferences and during lectures. The use of graphs and text allows speakers to convey the main points of their study to their audience. In particular, the tables and graphs from a data report, together with statistical software results, can be used in reports or in presentation slides. Adding a data visualization element to a presentation is a way to engage the audience. Copying and pasting data results to the presentation provides a quick, easy way to summarize a large amount of information in an understandable format. PivotCharts are especially useful for summarizing large amounts of complex data. Attendees at a presentation can get PDF copies of the presentations for later review.

Recommending Organizational Action

An inventory of all the knowledge repositories to record the name, location, user base, administrator, and content of the repository can provide a useful index to the current data stores. For example, PubMed may be used as a knowledge repository. The content should be specified down to the level of the data dictionary details of each data element in each repository. Data from the knowledge repository inventory should be placed in an enterprise-wide knowledge management metadata database so that it can be tracked and updated on a regular basis. This step is essential to understanding the data stored in each system.

Second, the report names, report content, when they are run, who requested the report, and how the results are disseminated should add to the metadata utilized by the healthcare facility. Recording the information about the facilities reports is important, as this can help top management identify issues with data integrity as well as issues with reports. Specifically, it is common for the queries that underlie these reports to summarize the detailed data in such a way that the summation misrepresents the actual data. This could result from misusing a database function, errors in the query code, or data type conversion issues (for example, number stored as character incorrectly converted). Decisions based on misrepresented data are likely to be incorrect.

Another frequent issue in complex reporting involving the same data element being stored in multiple locations is that the data are not updated in the same way in all locations. In the simplest case, this could occur when two departmental systems update the same field and the data then flow into the EHR. A second way that data integrity can be compromised occurs when data marts are used. To understand this, it is important to know that running large queries for reporting slows down a database. Moreover, an error

page 193

in the coding of the SQL reporting query can even cause the database processing to effectively be hung up with nothing being run. To facilitate reporting while not affecting database performance, a data mart is created for reporting. For example, the accounting department might want to see specific data from multiple tables in several databases, but they do not want to run 20 queries to do this. The DBAs will create a data mart (that is, set of tables) for use by the accounting department for reporting. The benefit here is that accounting users can efficiently run their reports with little or no access to the main database tables, but they cannot make any changes to the data. Of course, the downside is that the DBAs have to put data into these tables frequently or the data become old, and this creates the problem originally mentioned of updates running against several data marts not updating the data the same way or at the same time. In fact, if the lag in updating is great enough, one department could make uninformed decisions that flow into another department's data mart and wipe out the original data.

Case Study: Quantitative Study

Using surveys to collect and analyze data from a pool of respondents is a common quantitative research method. In Price et al.'s (2020) study, a national survey in the UK was undertaken to evaluate the impact of COVID-19 on neurosurgical oncology services and to explore changes in normal multidisciplinary team decision-making compared with the pre-COVID-19 era to investigate whether patients were denied access to surgery. Read this study and learn what information is collected in a national survey and how the collected data were analyzed to answer the research question.

Case Study: Electronic Health Records for Research

Since the Health Information Technology for Economic and Clinical Health (HITECH) Act was signed into law in 2009, most health care organizations in the US have adopted EHR systems. The volume of EHR data has increased dramatically in recent decades. The EHR has been used by researchers to conduct clinical and translational research and has been a major focus of health informatics research.

Chapter 7 Review Exercises

Match the research type with its description.

a. Quantitative
b. Qualitative research
c. Mixed methods
d. Retrospective

1. _____ Tests a theory by defining measurable variables and collecting and analyzing data on those variables
2. _____ Case control study examining past events

3. _____ Study with both open and closed questions
4. _____ Nonnumerical data are analyzed to investigate group interactions

Indicate whether the following statements are true or false.

a. When conducting research, the HIPAA privacy rule's minimum necessary standard must also be followed even when the use of the patient information is authorized or a waiver has been granted.
b. Research must be compliant with state laws and must be IRB approved.
c. The alternate hypothesis states that there is no difference between the independent and dependent variable.
d. The researcher should always check to see if the sample is similar to the population.

Discussion

5. A study on the correlation between age and hours of social media use per week is being conducted among physicians at your hospitals. Apply the steps in the research process to this study:
   - Determine the research question
   - State research hypotheses
   - Define study variables and their data types
   - Identify sample and population, data collection
   - Conduct data analysis

   For each step, state in one or two sentences what you would do at that step.
6. Discuss the role of the EHR in healthcare research and give at least two examples.
7. Select two statistical data analysis tools mentioned in the chapter that would be suitable for data mining on large amounts of quantitative data. Describe why you selected those tools with respect to their data collection, data analysis, and data graphing functionalities.
8. Select four data visualization tools mentioned in the chapter. Create a table listing their advantages and disadvantages.

REFERENCES

21 CFR Parts 50 and 56: U.S. Food and Drug Administration. FDA Policy for Protection of Human Subjects. 1991.

45 CFR 46.111(a)(7): Research—Protection of human subjects, including protecting confidential information. 2006.

45 CFR 164.502(b): Minimum necessary requirement. 2016.

45 CFR 164.514(d): Other requirements relating to uses and disclosures of protected health information. 2016.

AHIMA. 2017. ***Pocket Glossary of Health Information Management and Technology***. 5th ed. Chicago: AHIMA Press.

Aronson, A. R., J. G. Mork, A. Névéol, S. E. Shooshan, and D. Demner-Fushman. 2008. Methodology for creating UMLS content views appropriate for biomedical natural language processing. ***AMIA Annual Symposium Proceedings*** 2008:21–25.

Attia, Z. I., P. A. Noseworthy, F. Lopez-Jimenez, S. J. Asirvatham, A. J. Deshmukh, B. J. Gersh, R. E. Carter, X. Yao, A. A. Rabinstein, B. J. Erickson, S. and Kapa. 2019. An artificial intelligence-enabled ECG algorithm for the identification of patients with

atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. The Lancet, 394(10201): 861-867.

Barkhordari, M., M. Padyab, F. Hadaegh, F. Azizi, and M. Bozorgmanesh. 2016. Stata modules for calculating novel predictive performance indices for logistic models. ***International Journal of Endocrinology and Metabolism*** 14(1):1–5. doi:10.5812/ijem.26707.

Centers for Disease Control and Prevention (CDC). 2016. "EPI-Info." https://www.cdc.gov/epiinfo/index.html.

Centers for Medicare and Medicaid Services (CMS). 2016. "***Medicare Chronic Conditions Dashboard: Regional level."*** https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Dashboard/Chronic-Conditions-Region/CC_Region_Dashboard.html .

Cochran, G. L., L. Lander, M. Morien, D. E. Lomelin, C. Reker, and D. G. Klepser. 2015. Consumer opinions of health information exchange, e-prescribing, and personal health records. ***Perspectives in Health Information Management*** (Fall):1–12.

Comprehensive R Archive Network. 2023. www.cran.rstudio.com.

Creswell, J. W. 2011. ***Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research***. 4th ed. Upper Saddle River, NJ: Pearson.

Department of Health and Human Services (HHS). 2016. "Institutional review boards (IRBs)." http://www.hhs.gov/ohrp/.

Dheeraj, R., S. Xiaogang, P. A. Patrician, L. A. Loan, and M. S. McCarthy. 2015. Exploring factors associated with pressure ulcers: A data mining approach. ***International Journal of Nursing Studies*** 52(1):102–111. doi:10.1016/j.ijnurstu.2014.08.002.

Dolezel, D. 2015a. How to use relational databases: Data retrieval with Structured Query Language. ***Journal of AHIMA*** 86(11):22–27.

Dolezel, D. 2015b. Metadata offers roadmap to structured data. ***Journal of AHIMA*** 86(2):44–46.

Dolezel, D., and J. Moczygemba. 2015. Implementing EHRs: An exploratory study to examine current practices in migrating physician practice. ***Perspectives in Health Information Management*** Winter:1–15.

Downey, L. V., and L. S. Zun. 2010. The correlation between patient comprehension of their reason for hospital admission and overall patient satisfaction in the Emergency Department. ***Journal of the National Medical Association*** 102:637–643.

Eyre, H., A. B. Chapman, K. S. Peterson, J. Shi, P. R. Alba, M. M. Jones, T. L. Box, S. L. DuVall, and O. V. Patterson. 2021. Launching into clinical space with medspaCy: A new clinical text processing toolkit in Python. ***AMIA Annual Symposium Proceedings*** 2021:438–447.

Forrestal, E. 2016. Research Methods. Chapter 19 in ***Health Information Management Concepts Principles, and Practices,*** 5th ed., edited by P. Oachs and A. Watters. Chicago: AHIMA Press.

Green, S. B., and N. J. Salkind. 2013. ***Using SPSS for Windows and Macintosh***, 6th ed. Upper Saddle River, NJ: Pearson.

Haenke Just, B., D. Marc, M. Munns, and R. Sandefer. 2016. Why patient matching is a challenge: Research on Master Patient Index (MPI) data discrepancies in key identifying fields. ***Perspectives in Health Information Management*** (Spring):1–20.

Harvey, L. A. 2018. REDCap: Web-based software for all types of data storage and collection. ***Spinal Cord*** 56:625.

Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009. 42 USC sec. 139w-4(0)(2).

Health Resources and Services Administration (HRSA). 2016. "***How do we measure and improve clinical processes and outcomes?"*** http://www.hrsa.gov/healthit/toolbox/ruralhealthittoolbox/patientquality/measureprocesses.html.

Hoyt, R., S. Linnville, S. Thaler, and J. Moore. 2016. Digital family history data mining with neural networks: A pilot study. ***Perspectives in Health Information Management*** (Winter):1–14.

Huck, S. W. 2012. ***Reading Statistics and Research***. Boston: Pearson Education.

IBM Corp. 2023. ***IBM SPSS Statistics for Windows*** [Computer software]. Armonk, NY: IBM Corp.

Jackson, J. 2014. R programming language gaining ground on traditional statistics packages. ***PC World*** (August 20).

Jakobsson, N., and M. Svensson. 2016. The effect of copayments on primary care utilization: Results from a quasi-experiment. ***Applied Economics*** 48(39): 3752–3762. doi: 10.1080/00036846.2016.1145346.

Layman, E. J., and V. J. Watzlaf. 2017. ***Health Informatics Research Methods: Principles and Practice.*** Chicago: AHIMA Press.

Linux Foundation. 2022. PyTorch. https://pytorch.org/.

Liu, H., S. J. Bielinski, S. Sohn, S. Murphy, K. B. Wagholikar, S. R. Jonnalagadda, K. E. Ravikumar, S. T. Wu, I. J. Kullo, and C. G. Chute. 2013. An information extraction framework for cohort identification using electronic health records. ***AMIA Summits Translational Science Proceedings***:149–153.

Luther, S. L., J. A. McCart, D. J. Berndt, B. Hahm, D. Finch, J. Jarman, P. R. Foulis, W. A. Lapcevic, R. R. Campbell, R. I. Shorr, K. M. Valencia, and G. Powell-Cope. 2015. Improving identification of fall-related injuries in ambulatory care using statistical text mining. ***American Journal of Public Health*** 105(6):1168–1173. doi: 10.2105/AJPH.2014.302440.

Microsoft. 2023. Excel [Computer software]. Redmond, WA: Microsoft.

Minsker, M. 2015. Peek into the future. ***CRM Magazine*** 19(4):22–26.

Ostrovsky, A., L. O'Connor, O. Marshall, A. Angelo, K. Barrett, E. Majeski, M. Handrus, and J. Levy. 2016. Predicting 30- to 120-day readmission risk among Medicare fee-for-service patients using nonmedical workers and mobile technology. ***Perspectives in Health Information Management*** (Winter):1–20.

Patena, K. 2016. Employee Training and Development. Chapter 24 in ***Health Information Management Concepts Principles, and Practices,*** 5th ed, edited by P. Oachs and A. Watters. Chicago: AHIMA Press.

Potter, A. J., M. M. Ward, N. Natafgi, F. Ullrich, A. C. McKinney, A. L. Bell, and K .J. Mueller. 2016. Perceptions of the benefits of telemedicine in rural communities. ***Perspectives in Health Information Management*** (Summer):1–13.

Price, S. J., A. Joannides, P. Plaha, F. T. Afshari, E. Albanese, N. U. Barua, H. W. Chan, G. Critchley, T. Flannery, D. M. Foutain, and R. K. Mathew. 2020. Impact of COVID-19 pandemic on surgical neuro-oncology multi-disciplinary team decision making: A national survey (COVID-CNSMDT Study). ***BMJ Open*** 10(8):e040898.

R Development Core Team. 2023. ***R: A Language and Environment for Statistical Computing***. Vienna, Austria: The R Foundation for Statistical Computing. http://www.gbif.org/resource/81287.

Ramsey, K.S. ***Using Predictive and Descriptive Models to Improve Nurse Staff Planning and Scheduling*** [Master's thesis]. University of Tennessee, 2014 trace.tennessee.edu/cgi/viewcontent.cgi?article=4033&context=utk_gradthes.

RStudio. 2023. ***Welcome to R Studio.*** https://www.rstudio.com/.

Sandefer, R. H., D. T. Marc, and P. Kleeberg. 2015. Meaningful use attestations among US hospitals: The growing rural–urban divide. ***Perspectives in Health Information Management*** (Spring):1–10.

SAS Institute, Inc. 2023. ***Statistical Analysis Software (Version 9.4)***. Cary, NC: SAS Institute, Inc. http://www.sas.com/en_us/industry/health-care-providers.html.

Savova, G. K., J. J. Masanz, P. V. Ogren, J. Zheng, S. Sohn, K. C. Kipper-Schuler, and C. G. Chute. 2010. Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): Architecture, component evaluation and applications. ***Journal of American Medical Informatics Association*** 17(5):507–513.

Soyiri, I. N., and D. D. Reidpath. 2013. An overview of health forecasting. ***Environmental Health and Preventive Medicine*** 18(1):1–9. doi: 10.1007/s12199-012-0294-6.

Soyiri, I. N., D. D. Reidpath, and C. Sarran. 2013. Forecasting peak asthma admissions in London: An application of quantile regression models. ***International Journal of Biometeorology*** 57(4):569–578. doi:10.1007/s00484-012-0584-0.

Soysal, E., J. Wang, M. Jiang, Y. Wu, S. Pakhomov, H. Liu, and H. Xu. 2018. CLAMP—A toolkit for efficiently building customized clinical natural language processing pipelines. ***Journal of the American Medical Informatics Association*** 25(3):331–336.

SPSS. 2016. IBM Analytics. http://www.ibm.com/analytics/us/en/technology/spss/.

StataCorp. 2016. "***Resources for learning Stata.***" http://www.stata.com/links/resources-for-learning-stata/.

TensorFlow. 2022. https://www.tensorflow.org/.

Wang, Y., L. Wang, M. Rastegar-Mojarad, S. Moon, F. Shen, N. Afzal, S. Liu, Y. Zeng, S. Mehrabi, S. Sohn, and H. Liu. 2018. Clinical information extraction applications: A literature review. ***Journal of Biomedical Informatics*** 77:34–49.

White, S. 2016. Healthcare Data Analytics. Chapter 17 in ***Health Information Management Concepts Principles, and Practices,*** 5th ed., edited by P. Oachs and A. Watters. Chicago: AHIMA Press.

ADDITIONAL RESOURCES

45 CFR 160.502: Uses and disclosures of protected health information: General rules. 2016.

Absten, J. 2016. Tools and technologies. ***Industrial Engineer*** 48(2):56–57.

ComputerWorld. 2016. "Learning R for beginners with our PDF." http://www.computerworld.com/article/2884322/learn-r-programming-basics-with-our-pdf.html.

ITBusinessEdge. 2015. "SQLCourse—Lesson 3: Selecting data." www.sqlcourse.com/select.html.

Lau, F. Y. 2004. Toward a conceptual knowledge management framework in health. ***Perspectives in Health Information Management*** 1:8.

National Library of Medicine. 2010. "Comparative effectiveness research (CER)." https://www.nlm.nih.gov/hsrinfo/cer.html.

Sheridan, P. T., V. Watzlaf, and L. A. Fox. 2016. Health information management leaders and the practice of leadership through the lens of Bowen theory. ***Perspectives in Health Information Management*** (Spring):1–36.

The Complete ggplot2 Tutorial. http://r-statistics.co/Complete-Ggplot2-Tutorial-Part1-With-R-Code.html

Wickham, H. and G. Grolemund. 2016. R for data science: import, tidy, transform, visualize, and model data. "O'Reilly Media, Inc.".

Chapter

8

Introduction to Data Science

By Leming Zhou, PhD

Learning Objectives

- Describe the types of data, as well as the relationship between data and information
- Compare and contrast the methods of data collection
- Demonstrate the use of data quality practices
- Design appropriate data presentations
- Distinguish between unstructured data and structured data
- Articulate big data utilization in healthcare

KEY TERMS

All of Us Research Program
Big data
Categorical data
Continuous data
Crosstabs
Data
Data cleaning
Data science
Dependent variable
Descriptive statistics
Frequency
Imputation
Independent variable
Interval data
Natural language processing (NLP)

Nominal data
Ordinal data
Primary data use
Qualitative data
Quantitative data
Ratio data
Raw data
Regression
Secondary data use
Statistical modeling
Structured data
Unstructured data

Data and information are the resources that fuel the healthcare industry. Getting the right data or information to the right person at the right time is essential for the delivery of high-quality care as well as the financial well-being of healthcare provider organizations. Data can be entered into electronic health records (EHRs) to be used for clinical care, disease management, or disease prevention. They are also found in large data sets or data warehouses where clinicians, researchers, payers, and others access the data for a variety of purposes.

Data science is an interdisciplinary field. The focus of data science is on extracting information from data sets and applying the information for decision-making. To achieve these goals, skills from computer science, information science, statistics, mathematics, social science, and domain knowledge about the specific data sets are incorporated in these data to information and then to insight and decision process. The obtained solution is data-driven and evidence-based.

To use data science in healthcare, it is imperative that health informatics (HI) professionals understand the basics of data, issues involved with data collection, and the analysis and presentation of data.

The Basics of Data

Data are the "dates, numbers, images, symbols, letters, and words that represent basic facts and observations about people, processes, measurements, and conditions" (AHIMA 2017). Examples of data can include the entire set of entries for a single patient in the EHR or data sets such as the National Center for Health Statistics (NCHS), National Hospital Discharge Data Set (NHDDS), and the claims data of the Centers for Medicare and Medicaid Services (CMS). These data sets store the facts and observations about patients and the general population, which are invaluable for further analysis.

Data Related to Information

Although data are important, it is often difficult to utilize them effectively even for a single patient unless they are processed into information. So, how does the HI professional need from individual data elements to information that can be used effectively and efficiently? Foremost is an understanding of the original purpose for which the data are used or collected. For example, data that are collected for clinical purposes may be very different from data that are collected for reimbursement. Clinical data are often very detailed, for instance, vital signs can be collected every 15 minutes in

the intensive care unit (ICU), while the reimbursement data for the same patient will only include the more general diagnostic code.

It is important to consider whether there are mandated or generally accepted data standards for the data. Data standards may be related to the data set, such as the data elements specified for the Uniform Hospital Discharge Data Set, or the allowable data values, such as a selection list being provided for the address data element of State. Standards such as the allowable data elements within a data set, the definitions of the different data elements, or the allowable data values, which constrain the data that can be entered into a field, ensure that the data are as accurate as possible. Another data standard is the Value Set Authority Center (VSAC) from the National Library of Medicine (NLM), in collaboration with the Office of the National Coordinator for Health Information Technology (ONC) and the CMS. The VSAC provides data on the 2014 Clinical Quality Measures (NLM 2022).

An additional consideration is the data collection method that was used. There are many different methods of data collection, such as keyboard input, voice recognition, and automated methods, where data are transferred directly from digital thermometers, scales, blood pressure cuffs, and so on. The quality of these data items can vary widely.

The quality of the data must be assessed. It is difficult to produce good-quality information when the data are of poor quality to start. It is also vital to understand the analysis that the data are intended to undergo or the analysis that was used to develop or produce the information.

Finally, it is important to present the data so that they can be easily understood. Two commonly used forms are tabular and graphical presentations. It is generally easier to detect the change of values in one variable over time using the graphical representation. Figure 8.1 presents these issues in the form of questions HI professionals can pose while they are working with data.

These and other concerns are important to address to ensure the data and information used in healthcare can be relied on for all of the varied purposes required. This chapter examines these different aspects of data that are important to the production of information.

Data Format

Data collected for an EHR starts the data format life cycle as raw data. **Raw data** are unformatted combinations of numbers, pictures, audios, videos, text, symbols, and words. Data on patient names, diagnoses, admission, and healthcare payer typed into an Excel workbook would be considered raw data. Data that the user entered into a web page for electronic transmission to the EHR database is another example of raw data.

Figure 8.1. Understanding data and information

Questions to Ask about the Data or Information

1. For which purpose or use was the data originally collected?
2. Are there any data standards related to the data?
3. How was the data collected?
4. What is the quality of the data?
5. How will the data be analyzed to produce information?
6. How can the data or information be presented so that it is meaningful to users?

An example of raw data would be a comma delimited text string representing a patient's first name, last name, middle initial, insurance, and date of admission. For example, comma delimited text for these fields might be: "Green, Sue, A., Blue Cross, 02/12/15". These data can be parsed and stored to a database.

Once all the data have been added to the database, it is common to run various reports to assess the number of admissions, diagnoses, payers, length of stay, and charges. The outputs of these reports provide useful information to healthcare managers, planners, coders, and researchers. Data that are included in these reports have been processed for meaningful use and are now in the form of information available for decision-making.

Primary and Secondary Uses of Healthcare Data

There are two main ways to use healthcare data. The most common one is the **primary data use** for treatment or care of patients. The data are compiled as the care is being provided, such as in an outpatient visit to a physician or physical therapist or during an inpatient stay in a hospital. These data include the information that is compiled for assessment (report of physical exam, lab and x-ray results, vital signs) or treatment (physician orders, documentation of operative procedures, medication records). This is the clinical information that is used for care at the time provided and for subsequent medical care of patients. The other use is the **secondary data use**. Secondary use of health data is the "non-direct care use of personal health information (PHI)" (Safran et al. 2007). One example is performing research studies on the data to determine the risk factors for 30-day readmission after hospital discharge.

Public health reporting, reimbursement, and quality improvement are typical purposes for collecting healthcare data. The information extracted from these data can be very useful for healthcare providers, patients, and the general population. For example, one can find trends in cases of infectious diseases such as COVID-19 and monkeypox, and determine the best potential solutions to improve healthcare quality.

There are many other purposes for collecting healthcare data that are less known to the general public, such as provider certification or accreditation, disease management, determining best practices or clinical guidelines, detecting financial fraud, monitoring patient compliance, and identifying candidates using EHR data for clinical trials. In fact, "Collect once, use many" is a common phrase used to describe healthcare data that are collected for patient care purposes and then used to fulfill a myriad of additional requirements. Organizations describe "big data" to convey the message that the widespread implementation of health information technology (HIT) has resulted in extraordinarily large and detailed databases being made available for secondary data use.

Big Data

Big data are defined as "high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making" (Gartner 2013).

The first differentiator for big data is the sheer size, or volume, of the databases. Large hospital systems are accumulating terabytes of data every year. Likewise, the CMS databases from 1965 to present would qualify as big data.

The second issue is the speed, or velocity, with which the data need to, and do, move around the system. This pace will continue to accelerate as health information exchange organizations (HIOs) become more prevalent and are more widely used.

The third characteristic of big data is the variety of the data. The industry is beginning to seek ways to combine individually identifiable patient record information with genomic data, staffing data, wearable sensor data, mobile app data, or other types of data, which adds to the data's complexity. In particular, it might be difficult to find tools that can handle such a wide variety of data. Big data are and will continue to be a focus for HI

professionals because they need to develop new methods and tools to extract information from these massive data sets.

Two other characteristics of big data are not discussed as often as three mentioned ones. The fourth characteristic is veracity; that is, the data are very messy. The possible reason could be the poor quality of some data items or missing data. However, because there is a huge volume of data, some missing data and some data of poor quality might not be important issues in certain cases.

The fifth characteristic of big data is value; that is, one should be able to extract information and knowledge from big data. Otherwise, a huge data set will not mean anything to users.

Data Collection

Data can be collected via keyboard input, handwriting, speech recognition, or direct feed from medical devices, wearable sensors, and mobile health apps. Data collection should be carefully planned, keeping in mind the expected uses of the data and the needs and workflow of the users. Generally, data can be classified as either structured or unstructured. Figure 8.2 outlines questions that should be answered by organizations planning their EHR data collection efforts. Ultimately, the choices are usually between structured data, unstructured data, or a carefully considered combination of these two.

Structured Data

The structured entry of healthcare data involves the use of templates and on-screen forms with the possible entry fields with the potential entries in those fields controlled, defined, and limited (for instance, date selection on a provided calendar or a drop-down list). **Structured data** can have many benefits including completeness, quality (Johnson et al. 2008), and accessibility of the data for a variety of purposes (van Mulligen et al. 1998). Structured data are appropriate and highly recommended for data entry when the options are limited or are required to conform to a specific standard. For example, the collection of race and ethnicity data are specified by the US government (US Census Bureau 2020). Although this could be a free-text field, the need to report data in conformance with the government standard has led most HI professionals to limit the entries in the race and ethnicity fields. Structured data are often also highly desired for uses such as statistical analysis, clinical decision support, and quality measure reporting. For example, the International Classification of Disease (ICD) system was originally developed by the World Health Organization (WHO) for mortality reporting, or reporting the cause of death on death certificates (CDC 2022). The mortality data stored in the National Vital Statistics System (NVSS) can be used to present the characteristics of those who died in the United States, determine life expectancy, and compare mortality trends with other countries, which can help decision makers to make corresponding policies and realize improvements in this area.

Figure 8.2. Data collection considerations

- Why is the organization entering these data? Is it for patient care? To secure payment? To comply with a regulation?
- What data need to be entered? Is there a standard for the data?
- Who will enter the data? Will it be clinicians, nurses, ancillary staff, clerks, or a combination?
- What will be the most efficient method of data entry for each type of user?
- Under what circumstances will users enter data? Will they be rushed, or will they have adequate time? What will be the setting (admissions, emergency room, intensive care, or the patient's home)?
- What use does the organization expect to make of the data? Will they be used for patient care, reimbursement, quality improvement, or research?

Source: Fenton (2006)

Structured data are not without their negative aspects. Some studies have suggested that structured data entry can slow down users, decreasing their productivity (Ash et al. 2007; Rosenbloom et al. 2011). In addition, structured data entry tools can be incomplete for user needs as well as inflexible when new or unexpected demands arise (Rosenbloom et al. 2011).

Therefore, it is important to continue to allow and even encourage the use of unstructured data in healthcare. Unstructured data provide details. HI professionals must develop methods and processes to convert unstructured data to structured data for secondary use purposes. Table 8.1 provides some specific examples of structured and unstructured data.

Unstructured Data

Unstructured data does not mean that there is no organization to the data, just that the data are not structured with specific data elements, allowable values, specific sizes of the data, and so on. Unstructured data might be organized into data fields or text boxes such as the history of illness and a progress note for the physician's hospital visit. However, this organization only conveys very general meaning and is usually only relevant regarding the type of documentation one can expect to find. Inferring meaning from native, unstructured data is difficult. Unstructured data require more complex processing before use as secondary data, but they provide a rich store of patient information. Unstructured data are often preferred because they enable providers to document details and nuances that are usually not available with structured data. For instance, it is very hard to use structured data to capture all types of patient symptoms and a doctor's impression after a physical examination, but these data can be easily captured with unstructured clinical notes. In healthcare documentation, unstructured data include most provider notes during a course of care, the history of an illness, the description of a procedure, or any situation where the number of different permutations is almost infinite. After all, each patient is unique and doctors need the flexibility to document that uniqueness.

Table 8.1 . Examples of structured and unstructured data

Structured data	Unstructured data
Data entered into templates Coded data (for example, ICD, CPT, SNOMED) Bar codes	Progress notes Video and audio files Scanned documents CT, x-ray, and MRI images

Unstructured data can be further classified by whether or not they are textual or nontextual. Textual unstructured data are based on text such as e-mails and other free-form entries or documents. Nontextual data can be digitized images, audios, and videos. Picture archiving and communications systems (PACS) store digitized medical images, such as ultrasound or radiology images, and electronically transfer the medical image data in a Digital Image and Communications in Medicine (DICOM) format.

One challenge with unstructured data usage is the ongoing and increasing need to use the data in ways that require them to be structured. If the data are textual, the solution is **natural language processing (NLP),** which is a branch of artificial intelligence that uses computer algorithms and statistical probabilities to convert human free text into structured data formats. If the data are images, audio, or video files, the current approach is deep learning, which is one type of machine learning approach. Deep learning is particularly effective in identifying patterns, determining differences, and recognizing objects in those nontextual files.

Whether an organization chooses structured or unstructured data entry depends on the purpose of the data entry and the eventual uses of the data. Data entry acceptable for administrative purposes might not be adequate for clinical processes. Organizations must carefully investigate data entry when planning data collection. Ultimately, each organization will need to make its own carefully considered decision when selecting data entry methods for its own data collection (Fenton 2006).

Metadata

To use data and information appropriately, it is important to understand metadata. Metadata are

> descriptive data that characterize other data to create a clearer understanding of their meaning and to achieve greater reliability and quality of information. Metadata consist of both indexing terms and attributes. They are data about data: for example, creation date, date sent, date received, last access date, last modification date. (AHIMA 2017)

Metadata are essential for a full understanding of the underlying data. For example, without metadata such as the creation or modification date, it might be impossible to determine which version of ICD-10-CM was used for coding a record. This is essential to ensure the code is interpreted correctly. Metadata can help explain the data or information by defining the data element or specifying the length of the field or providing the values that are allowed to be entered in the field. Metadata can change over time. For example, a healthcare organization might have specified metadata related to its EHR data; however, if the EHR data are transmitted, a new metadata field could be added to indicate the healthcare organization as the source of the data. Metadata are one of those hidden, necessary tools for the effective use of data and information.

Check Your Understanding 8.1

1. Primary data use is required for which of the following?
   a. Direct patient care
   b. Public health reporting
   c. Provider certification or accreditation
   d. Monitoring patient compliance
2. Which type of data provide details?
   a. Natural language processing
   b. Unstructured data
   c. Primary use data
   d. Secondary data
3. Which of the following would be considered metadata?
   a. ICD-10-CM codes
   b. Source of the data
   c. Trend analysis
   d. Data of admission
4. Which is not an example of the secondary use(s) of data?
   a. For accreditation purposes
   b. Determining best practices
   c. Quality improvement activities
   d. Review by a consulting physician

Indicate whether each of the following is true or false.

5. Unstructured data are generally more difficult to use than structured data.
6. Physician progress notes would be considered structural data.
7. Data without metadata are just as easy to understand as data with metadata.

Indicate which type of data entry (s = structured; u = unstructured) applies to the pros and cons.

Pros

1. _____ The data will have detail and nuance.
2. _____ The data are easily used for analysis and reporting.
3. _____ It is easy to determine if these data are complete.
4. _____ These data usually reflect more of the clinician's thought processes.

Cons

5. _____ It usually takes more time to enter data using this method.
6. _____ It is more difficult to use these data for reporting or clinical decision support algorithms.
7. _____ This type of data entry tends to be very inflexible.
8. _____ It is difficult to determine the completeness of these data.

Qualitative and Quantitative Data

The selection of the correct measurement scale to measure data is very important because it determines the type of statistical tests that can be conducted. The four scales of measurement in common use are nominal, ordinal, interval, and ratio. These four measurement scales represent two types of data, qualitative and quantitative, or categorical and continuous. **Qualitative or categorical data**, which include nominal and ordinal scale data, represent mutually exclusive categories or labels, such as race, sex,

ethnicity, country, state, and so on. **Quantitative or continuous data** such as interval or ratio scale data have equal intervals between the data points. **Measurements of continuous data can contain fractions and decimals in the data.**

Nominal Data

Nominal data are a type of qualitative data where the categories are simply names. Demographic data that are collected on a survey using check boxes are often nominal data; for example, marital status with choices such as married, single, divorced, and such. Other nominal data frequently collected are age group, race or ethnicity, type of insurance payer, and job category. Data collected for two patients on a nominal scale can be determined to be different, but the data analyst cannot add or subtract the data to compute the amount of difference because nominal data are not numerical. For that reason, the summary of nominal data is typically reported using percentages or frequency counts.

Ordinal Data

Ordinal data, also called ranked data, are qualitative data where the names or labels have an order to them with meaning attached. A Likert rating scale on which respondents select, for example, strongly agree (1), agree (2), neither agree or disagree (3), disagree (4), and strongly disagree (5), is an example of ordinal data. Likert scales are used for staging cancer, which is often represented as Stage I, Stage II, Stage III, or Stage IV. These are mutually exclusive categories; however, there is an order to the labels, with Stage IV cancer being more advanced than Stage I cancer. Data represented in an ordinal scale can be identified as different and a direction can be associated with that difference (higher or lower on the scale). Percentages and frequency counts can also be used to summarize ordinal data.

Interval Data

Interval data are continuous numerical data that have equal intervals of ordered categories. These data are called continuous because there is an infinite number of possible data measurement values. Because the intervals between the data measurements are equal, these measurements can be compared to determine meaningful differences. Temperature in degrees Fahrenheit is an example of interval data. Note the interval data have an ***arbitrary*** zero value, which does not indicate an absence of that value. For instance, 0 degree in Fahrenheit does not mean there is no temperature. This type of data can be compared to quantify the differences in two interval measurements.

Ratio Data

Ratio data are a type of continuous numerical data. In addition to the equal intervals between data points that are found in interval data, ratio data have a true zero (0) point. An example of ratio data in healthcare would be weight. No human can weigh less than 0 pounds or ounces so this data element would be ratio. Data on age, length of stay, and charges are measured on a ratio data measurement scale.

Check Your Understanding 8.2

Match the terms with the appropriate descriptions.

a. Nominal
b. Ordinal
c. Ratio

d. Interval

1. _____ These data do not have a true zero and the distance between the data points is consistent.
2. _____ These data separate records into mutually exclusive categories in a meaningful order.
3. _____ These data have a true zero and the distance between the data points is consistent.
4. _____ These data separate records into mutually exclusive, named, categories with no order.

Answer the following questions.

5. Which data would be found in a Likert scale on a survey?
   a. Ordinal
   b. Interval
   c. Nominal
   d. Ratio
6. Temperature in degrees Fahrenheit is an example of this data type.
   a. Nominal
   b. Ordinal
   c. Interval
   d. Ratio
7. This type of data typically has many measurements.
   a. Ordinal
   b. Continuous
   c. Discrete
   d. Nominal

Preparation for Data Analysis

Preprocessing

The increasing computerization of the healthcare industry has left many organizations awash in data. Effective data analysis plans, techniques, and processes are needed to turn the data into usable information. That typically starts with data preprocessing, which is a procedure used to manipulate and cleaning raw data so that they are ready for further analysis, such as descriptive statistics. Typical data preprocessing steps include identifying and removing errors in the data, handling missing data, and organizing data according to a chosen standard.

Understanding the Data

The first and most essential step to analyzing any data is acquiring a thorough understanding of the data to be analyzed. This is best done by studying the **data dictionary, if available.** The data dictionary is vital to understanding the data set.

The data dictionary might include information related to the originating source system, data owner, data entry date, and data termination date (AHIMA 2017). In addition, the data dictionary should include information regarding special processing uses in the production of the data set. It is common for some data records to have missing values.

One part of understanding the data that can easily be overlooked is obtaining an understanding of the data collection methods. Laboratory data that are loaded directly into a laboratory information system and then transferred to an EHR automatically are

likely to be of better quality than data that were entered manually. Each time the data are handled or touched by a human the opportunity for or likelihood of error increases. However, this understanding of data collection can also be subtler. Consider many claims databases that are used for research, quality measure development, and so on. Although these are often the most easily accessed, they might not always be the best. In ICD-10, the external cause codes (E-codes) were replaced by V, W, X, and Y codes that can be used for the entire treatment (Endicott 2014). Other hospitals code only cover the principal diagnosis and the secondary diagnoses, which will affect reimbursement. Often these different methods of data collection are made for very pragmatic reasons such as coder workforce shortages. However, a failure to understand these differences when using large data sets from multiple organizations could result in false conclusions being drawn regarding the causes of injuries or the presence or absence of certain health factors.

Cleaning the Data

Many data analysis projects begin with cleaning the data. **Data cleaning** involves examining the data thoroughly to detect wrong or inconsistent data. This is a necessary task even if the data are drawn from a data set that is managed with a data quality program. When data are subset or transferred between data applications, the opportunity exists for the introduction of errors.

The first step in data cleaning is to load and label the data. This process could be largely automated or data analysts might have to perform it themselves. Either way, the data analysts need to validate the data elements in the data set against the data dictionary.

Once the data are loaded and labeled, each data element needs to be examined. **Descriptive statistics,** also known as summary statistics, such as frequencies, mean, median, range, and standard deviation, should be run for all continuous data elements. Table 8.2 is an example of descriptive statistical output for the continuous data element of age. On initial examination, the output seems fine, with a reasonable mean and median. However, further examination of the output reveals a range of 268. Without looking at the minimum and maximum, the data analyst can tell that errors exist for this data element because 268 is not a valid entry for age when the unit is years. Most likely at least one incorrect, large entry was made for age. In this instance the minimum is 23.3, which is entirely reasonable, but the maximum is 291.3, which is not reasonable.

Correcting the continuous data element error first necessitates identifying the record(s) with the error. If available, it is always preferable to return to the source data to determine the correct value. If the source data are not available, the correct value may be imputed, or the incorrect record might have to be eliminated from the analysis.

Table 8.2 . Descriptive statistics for a continuous variable

Age	
Mean	60.834375
Standard Error	3.851114233
Median	66.4
Mode	N/A
Standard Deviation	21.78519191
Sample Variance	474.5945867
Kurtosis	–1.25981969
Skewness	–0.38286422
Range	268

Minimum	23.3
Maximum	291.3
Sum	2146.7
Count	32

Categorical data require a different method of examination for data cleaning purposes. Frequencies should be run on all categorical data elements. Table 8.3 is an example of a **frequency** for a categorical data element. A frequency is the number of times that a particular observation or value occurs in a data set. This type of output should prompt the data analyst to examine the data dictionary for the allowable values for this data element because "99" is not a typical discharge disposition. Either the label for a valid entry was not captured correctly or there are a significant number of records (50 out of 300) with an invalid entry.

The relationships between the data elements might also prompt categorical data cleaning using **crosstabs.** Crosstabs serve to highlight any errors where there is an expected or explicit relationship between two data elements. Table 8.4 is an example of a crosstab analysis for two categorical data elements. Because hysterectomy is a female-only procedure, the number of entries for male should be 0, although in table 8.4 it is shown as 5.

The use of crosstabs for data cleaning is more ambiguous than the use of descriptive statistics or frequencies for single data elements. Often the data cleaning using crosstabs may not be done until the type of data analysis being attempted is known.

Correction of categorical data errors can be handled in a manner similar to continuous data errors. First the record with the error must be identified. However, unlike continuous errors where the only options are imputation or reviewing the source data, the correct categorical data can sometimes be deduced from other data element entries. For example, the records with the hysterectomy errors in table 8.4 could be examined to attempt to determine whether the procedure was incorrect or the gender was incorrect. Other data that might provide a clue include preventive services such as a prostate exam or mammogram. The other options such as imputation and reviewing the source data are also available.

Table 8.3 . Frequency exploration of categorical data

Discharge Disposition	Count	Percent	Cumulative Percent
Home	100	33.33%	
Nursing Home	50	16.67%	50.00%
Rehabilitation	50	16.67%	66.67%
Died	50	16.67%	83.34%
99	50	16.67%	100.00%
			(rounding)
TOTAL	300	100%	

Table 8.4 . Crosstab of categorical data

Procedures by gender	Male	Female	Total
Laparoscopy	200	350	550
Cardiac Catheterization	40	25	65

Arthroscopy	350	300	650
Hysterectomy	5	25	30
TOTAL	595	700	1295

The most sophisticated analysis is useless if the data that went into it are not cleaned because the statistical calculations will be affected.

Imputation

When data are initially collected for a research study, the data set should be examined to determine if there are any missing data. Missing data values cause the data set to be incomplete. This must be addressed before statistical testing can be completed. Data could be missing due to errors such as data collectors not recording all the data in the data collection tool. Study participants might choose to not answer all the questions in a survey, which results in missing survey response data for those questions. When using a wearable sensor to collect data, study participants might take off the device at times. Another scenario that results in missing data occurs when the data are collected, but later misplaced or destroyed due to physical destruction or to a system failure that affects the storage medium (for example, a hard drive failure). Healthcare studies using instruments to collect data may experience instrument failure, such as the battery of a wearable sensor running out of power. Natural disasters could cause data loss from physical destruction of equipment or damage to a lab. Participants in a study might decide to leave the study or might have changed circumstances, making them unable to complete the study due to moving, scheduling conflicts, or other personal situations. For a long survey, respondents might get survey fatigue and start skipping questions, or perhaps they do not understand the questions.

The first step in addressing missing data is to determine what values are missing. For a small data set, the data can be examined by inspection (that is, looking at the data) to determine what is not present. In a larger Microsoft Excel spreadsheet, missing data would show up as blank cells, which can be located using the built-in Find and Select feature. For a very large data set, statistical software tools, such as Statistical Package for the Social Sciences (SPSS), or programming languages such as Python and R, should be used to locate any missing data, as they have a rich set of tools for this purpose. Graphing the data can assist in visualizing the data so that the location of missing data can be identified, and frequency tables will give the number of missing data items.

There are three common approaches to dealing with missing data. First, if question responses are missing for one participant, researchers might decide not to use any of that participant's data in the data analysis (Green and Salkind 2016). This creates a problem if there are a lot of missing data because there will be a loss of data for the study. Thus, data analysts could decide to include the participant's other responses in the study, in which case it should be noted in the study that the case was retained but excluded from some of the study calculations due to missing data for that response item (such as a survey question was skipped). This approach must be used judiciously because, for example, many participants might have skipped the same question due to a study design issue such as poor wording in the question or unclear directions. In this case the missing data are not randomly distributed and excluding them could skew the data (Cokluk and Kayri 2011).

Second, the missing data could be coded with an identifier such as -9999 to be more easily located and dealt with during the data analysis. For example, statistical software can be programmed to exclude this user-defined value from calculation.

Third, depending on the type of analysis, the missing values might be provided using a process called **imputation** in which missing data values are replaced with some

calculated values (Cokluk and Kayri 2011). Imputation methods must be described fully and imputed values clearly labeled. For numerical data, the mean of all other values for that data item can be substituted for all the missing values. Other commonly used replacement values are simple calculated values such as the mean or median of points near the missing value, and complex predictive values calculated with regression. It should be noted that replacing too much data could also input bias into the study. All procedures used to replace missing data must be documented.

Check Your Understanding 8.3

1. A data dictionary consists of
   a. The names, definitions, and attributes of data elements
   b. Synonyms for specific data variables
   c. All possible data elements in the healthcare industry
   d. Hints for using data elements
2. Once any records with data errors are identified, the data analyst should
   a. Delete the records with the errors
   b. Perform imputation to determine the correct values
   c. Determine whether the source records are available to validate the data
   d. Attempt to correct the data from the other data in the record

Indicate whether the following statements are true or false.

3. The first, most essential step in analyzing any data is acquiring a thorough understanding of the data to be analyzed.
4. Data cleaning is one of the smallest parts of a data analysis project.
5. Descriptive statistics include the frequencies, mean, and standard deviation.
6. A frequency is the number of times that a particular observation or value occurs in a data set.

Data Visualization

Healthcare activities generate large amounts of data from EHRs; computer physician order entry systems; digital radiology storage software; pharmacy, lab, and registration systems; and numerous other healthcare applications. Healthcare professionals struggle with finding ways to analyze this mountain of data into meaningful, actionable information.

Data visualization helps the data analyst to see the general shape of the data, which facilitates identifying skewed data, finding data outliers, exploring trends, and identifying possible missing data points. For example, a bar chart displaying customer complaint categories from highest to lowest is an efficient visualization of the upper 20 percent of the customers' complaints, which if resolved will theoretically reduce the number of complaints by 80 percent (Shaw and Carter 2015). Figure 8.3 displays a chart of complaint categories for a hypothetical clinic.

In statistical analysis, it is important to determine the distribution of the data before applying statistical methods to the data, for instance, whether the data are normally distributed. There are multiple ways to obtain an idea about the data distribution, such as histogram, density plot, and Q-Q plot. A histogram is an approximate representation of the distribution of numerical data and it is widely used during the data exploration stage of the data analysis. Figure 8.4 shows two histograms of body temperature measured over

a defined period of time. It is clear to see that both data sets do not have a normal distribution, but the distribution of temperature 1 is closer to a normal distribution than that of temperature 2. A more precise evaluation on data normality should be checked with statistical tests such as Shapiro-Wilks test, Kolmogorov-Smirnov test, skewness, and kurtosis (Mishra et al. 2019).

page 217

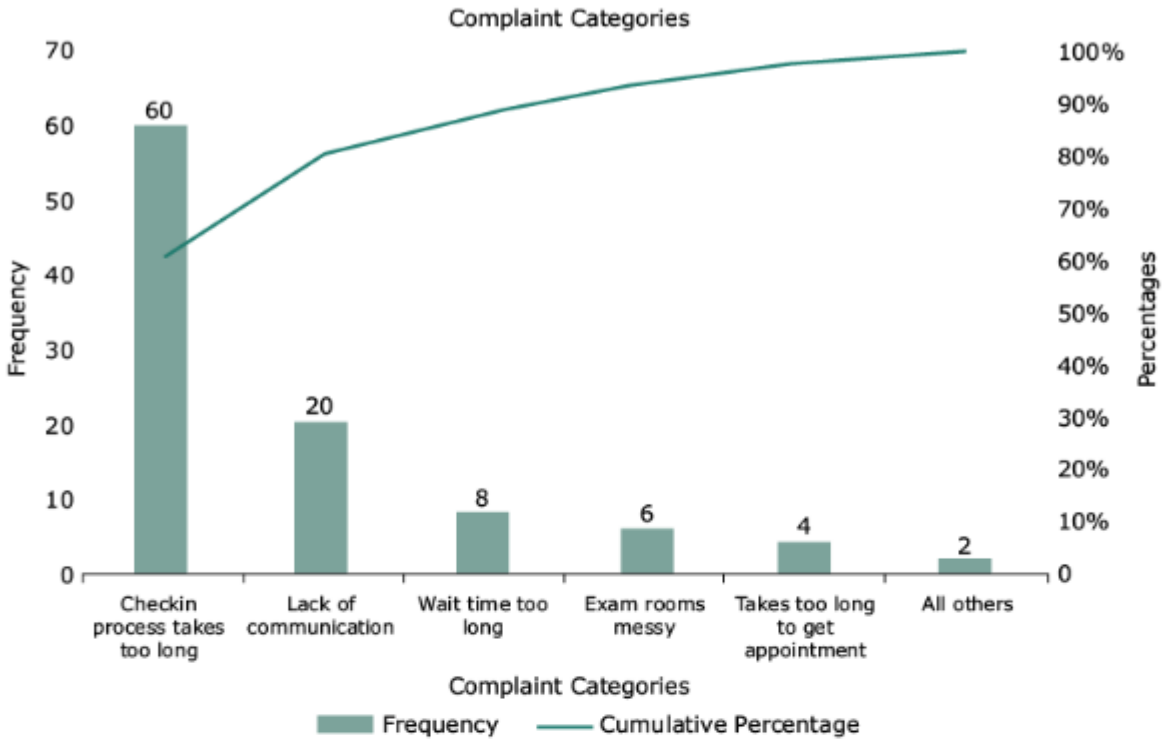


Figure 8.3. Bar chart of complaint categories for a clinic

▶ Expand/collapse Extended Description

In some cases, the given data are not numerical or categorical, but free text. We still have multiple ways to visualize the data to obtain a general idea about them. The first one is a word cloud diagram. In a word cloud, the importance (for example, the frequency of a word) of each word is shown with font size, color, and location. Important words are shown with larger font sizes, darker or brighter colors, and at the center of the diagram. Figure 8.5 demonstrates a word cloud generated from a list of paragraphs extracted from a few hundred journal articles. Just from this word cloud one can tell that the theme of these selected paragraphs is about data security because of the keywords such as data, secure, and encryption. Other important keywords, such as mobile, apps, server, database, and patient, can provide further information about the data. In fact, all the selected articles are journal papers about mobile health apps, and all of them have some descriptions about the information security technologies used to protect patient data, such as encryption, secure messaging, and access control. To obtain a more precise and quantitative sense about the keywords used in the text data, data analysts can also choose to explicitly calculate the word frequency and draw a bar chart for the top 20 keywords.

page 218

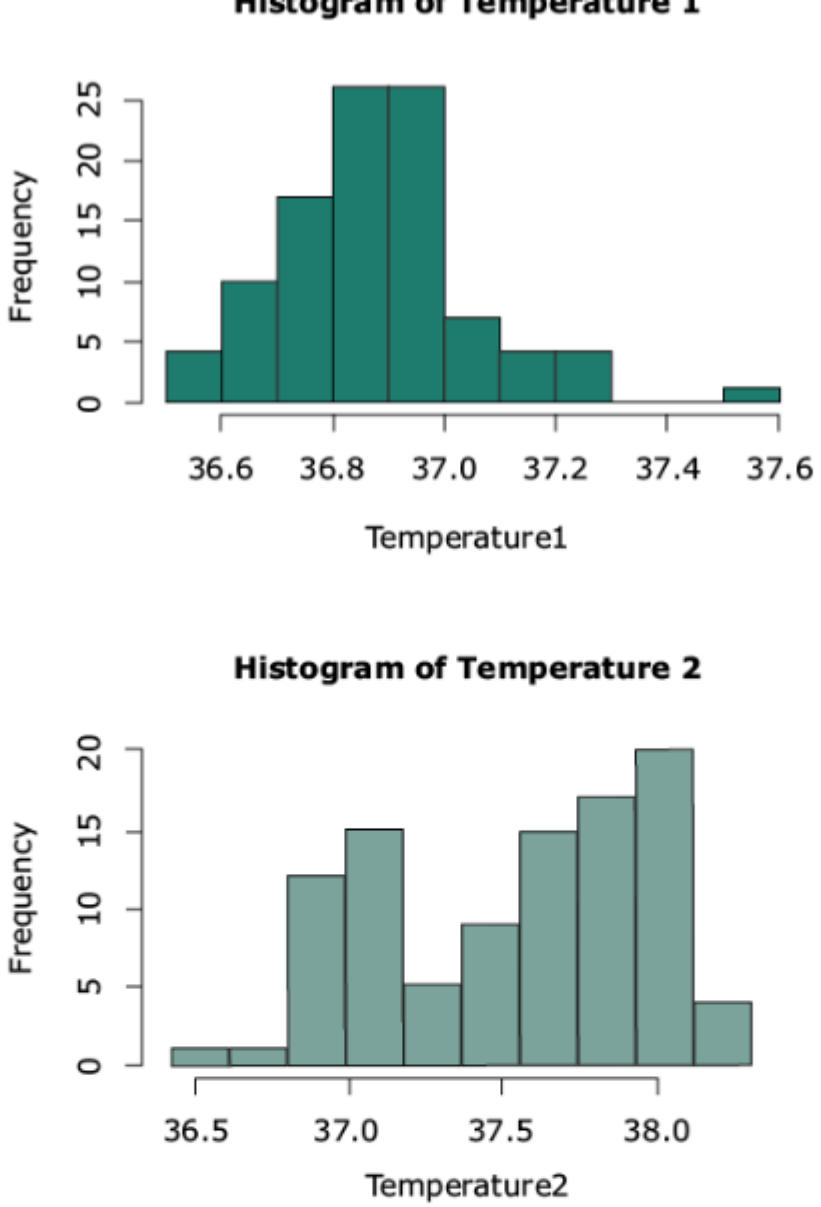


Figure 8.4. Histogram of body temperature measured in a certain period of time

▶ Expand/collapse Extended Description

Guiding Principles for Charts and Graphs

One of the benefits of a highly computerized world is that it is now relatively easy to generate charts and graphs for data exploration and analysis results. Charts and graphs can be very useful for emphasizing important points and conveying information about a topic. Guiding principles for the creation of charts and graphs include the following.

- Distortion—The representation of numbers or percentages should be proportional to the quantities represented.
- Proportion and scale—Graphs should emphasize the horizontal and be greater in length than height. A general rule is that the *y*-axis (height) be three-quarters the *x*-axis (length) of the graph.
- Abbreviations—Any abbreviations used in the chart should be spelled out in the caption for clarity.

- Color—Color should be used as appropriate to the use of the graph.
- Text—The font and use of capitalization needs to be considered carefully. The use of all capital letters can sometimes be difficult to read (Marc 2016, 542).

Ideally, spatial representations with graphs will

- Support the interpretation of the data
- Allow data trends to be seen at a glance
- Present individual data points clearly
- Present differences in data between groups or over time
- Avoid presenting too much data in too small a space (Marc 2016, 542)

It is important to determine the best type of graph or chart for data display. For example, a pie chart would not be a good choice for displaying data from two or more variables because a pie chart shows percentages for one variable only, such as percentages of physicians by specialty for the hospital. Data type is important when selecting a chart. For example, bar charts are used for nominal data and line charts can be used for displaying large amounts of data over time.

Figure 8.5. Word cloud generated from selected paragraphs in a few hundred journal articles

▶ Expand/collapse Extended Description

Data Visualization Applications

Data visualization can be applied to many fields and purposes. In one study, academic analytical and text mining software was used by researchers to model document metadata as part of a medical and healthcare curriculum mapping study (Komenda et al. 2015). In a related study, the spatial component of epidemiological data was explored for disease patterns using the HealthTracks web-based geographic information system software to graph demographic and healthcare data, such as community immunizations and notifiable diseases, for the purpose of identifying risk factors based on geolocation (Jardine et al. 2014).

In the personal injury area, Allscripts software has been used to provide timelines for body mass and body maps showing sites of previous injury. NodeXL software analyzed Twitter user data by searching for keywords like polio (Shneiderman et al. 2013). This type of data mining with large amounts of data can help guide medical researchers by uncovering new relationships in the data.

Another use of data visualization is to create data mapping. In a study on managing surgical outcomes, hierarchical mappings, called treemaps, of retrospective surgery data were evaluated as one way to help busy surgeons assess 30-day risk-adjusted surgical outcomes, surgical errors, and time to complete the surgery (Hugine et al. 2014). Results indicated that treemap visualizations were useful, but further research was needed into the optimal data displays for the data.

Data Analysis

Exploratory Data Analysis

Before choosing the statistical method or algorithm for analyzing a given data set, data analysts can perform some exploratory data analysis (EDA). At this stage, descriptive statistics can be used to check the distribution, summary of the data, correlation among variables, and similarities and differences among different samples. If the data set is highly complex with a huge number of variables (for example, >1,000), it could take quite a while to finish this step.

In SPSS, Python, and R, there are methods for obtaining descriptive statistics for all numerical variables in a loaded data set. For instance, in SPSS, options for descriptive statistics, such as mean, sum, standard deviation, variance, range, minimum, maximum, standard error mean, kurtosis, and skewness, can all be shown on one page. One simply needs to determine what descriptive statistics are needed for the specific analysis. In R, summary method can provide mean, median, 25th and 75th quartiles, minimum, and maximum. If the Hmisc package is installed, the describe method can provide more details, such as the number of data items; number of missing values; number of unique values; 5, 10, 25, 50, 75, 90, and 95th percentiles; and five lowest and five highest values. In Python, after data are loaded with the Pandas package, the describe method can provide a summary for each variable in the data set, such as the number of data items, mean, median, standard deviation, minimum, maximum, and 25th and 75th percentiles.

If the variables are not numerical, such as nominal and ordinal data, most of the summary statistics are not applicable, as there is no way to calculate mean, standard deviation, and variance on this type of data. For nominal data, one can determine the frequency of each category and identify the one with the highest frequency (mode). For ordinal data, one can do the same as on nominal data. Additionally, because the values of ordinal data have a specific order, it is also possible to determine the median of the data.

Correlation can be calculated between variables to determine whether two variables are highly correlated. In some studies, the results of one single lab test could be shown in two different columns with different units. When one performs data analysis or statistical modeling, only one of these two should be used. This issue can be detected by carefully reading the data dictionary if available or by calculating the correlation coefficient between these two variables. For instance, among the lab data in the National Health and Nutrition Examination Survey (NHANES), one of them is the measure for high-density lipoprotein (HDL; CDC 2020). There are three columns in this data file: SEQN, LBDHDD, and LBDHDDSI (Figure 8.6). The first column is the respondent sequence number, the second one is HDL-cholesterol in mg/dL and the third one is HDL-cholesterol in mmol/L. The second and the third columns have the same information. Because the units are different, the values in these two columns are different as well. If a data analyst did not read the data dictionary carefully, he or she might consider these to be two different variables. However, if he or she calculated the correlation coefficient, the result would be 1.0, which would clearly indicate the close relationship between these two variables.

	SEQN	LBDHDD	LBDHDDSI
0	93705.0	60.0	1.55
1	93706.0	47.0	1.22
2	93707.0	68.0	1.76
3	93708.0	88.0	2.28
4	93709.0	65.0	1.68
5	93711.0	72.0	1.86
6	93712.0	48.0	1.24
7	93713.0	48.0	1.24
8	93714.0	42.0	1.09
9	93715.0	57.0	1.47

Figure 8.6. First 10 rows of the lab data for HDL

In the previous section, when data visualization was discussed, the major focus was on individual variables. Data visualization can also be used during the EDA to show the relationship among multiple variables. For example, a boxplot can be used to show the minimum, maximum, median, 25th and 75th quartiles of three variables, A, B, and C (see Figure 8.7). The information could be used to guide the further analysis of these variables. For instance, if one wants to figure out whether the difference of means between two variables is statistically significant, one might want to start with a test on B and C, and then A and C. Similarly, a scatterplot can be very useful to show the relationship between two variables, which can be used to guide the method selection for regression. For instance, if the relationship between two variables is linear, one can use a linear regression to build a statistical model. If the relationship is not linear, a higher order regression is needed, or one variable might need to be transformed before performing the regression, for instance, applying a log or exp function on the variable.

Inferential Statistics

In the EDA, descriptive statistics can be used to provide a general idea about the individual variables and the relationship between variables. In the descriptive statistics, all the results are based on the data itself and the results are only about the data available in the sample. If one wants to draw conclusions about a population from the sample, inferential statistics are needed. For example, using descriptive statistics, we can easily calculate the means of two samples, A and B. In some cases, they are equal or have a small difference. However, the values of these two sample means alone cannot tell us whether the means of the two corresponding populations are equal or have a statistically significant difference. To find the desired answer for the two populations, one can perform hypothesis testing in inferential statistics.

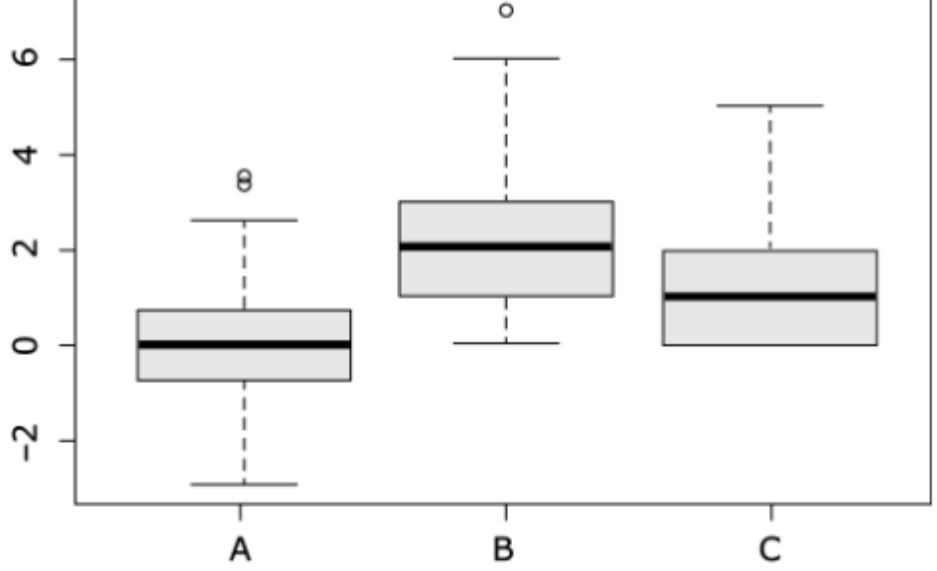


Figure 8.7. Boxplot for three variables

▶ Expand/collapse Extended Description

In hypothesis testing, there are two hypotheses; one is called the null hypothesis and the other is called the alternate hypothesis. For the example mentioned above, here are the statements for the null hypothesis and alternate hypothesis:

- Null hypothesis: There is no difference between the mean of A and the mean of B.
- Alternate hypothesis: There is a difference between the mean of A and the mean of B.

We can choose a statistical test according to the nature of the samples and conduct the hypothesis testing. The result will report some statistics and they are different for different statistical tests, but one of them would be a p value. Typically, 0.05 is used to determine statistical significance. If the p value is less than 0.05, the findings are unlikely to be due to chance (5 percent); therefore, we reject the null hypothesis. In other words, if p is smaller than 0.05, the conclusion is that there is a statistically significant difference between the population mean of A and B.

There are many different statistical methods for a hypothesis test. All of them are designed with specific assumptions and for different types of data. Data analysts need to choose the right one for the specific data. For example, if two samples have continuous data, and they are normally distributed, one might consider using a parametric test such as Student's t-test for the hypothesis testing. If the two samples are independent, an independent measure t-test should be used. If the two samples are the repeated measures of the same group of people in two different conditions, the paired sample t-test should be used. If the two samples do not have a normal distribution, a nonparametric test should be considered. If the two samples are independent, the Mann-Whitney test can be used; otherwise, a Wilcoxon test could be used instead for the hypothesis testing. If there are more than two samples and each is for one specific condition, a similar selection procedure should be used, but the specific tests are different.

Statistical Modeling

In many cases, it is well known that one intervention (for example, a medication) generates different outcomes in one population; however, it is not clear what specific factors (for example, age, gender, race, weight, smoking history, comorbidities, and such)

contribute to this difference and how big the contribution of each factor is. To find answers to these questions, one can build statistical models and run them on the collected data set. This is called **statistical modeling**.

Before one can perform the statistical modeling, two concepts must be clearly defined: dependent variable and independent variable. **Dependent variables** are targets, outcomes, responses, or class variables in a study. Their values are generated or determined from other variables. For example, when a group of patients with the same disease receive the same treatment, the outcome of the treatment can be better, worse, or have no change. This outcome is a dependent variable. **Independent variables** are predictors or explanatory variables and they determine the value of the dependent variable. These are the variables we provided examples of in the previous paragraph. The values of independent variables are collected of those patients and are not determined by other variables. In statistical modeling, one wants to find the contribution of each independent variable in the generation of the value of the dependent variable. For that purpose, a frequently used technique is **regression**.

There are a number of regression models, such as simple linear regression, general linear regression, and logistic regression. One can choose a model according to the nature of the data, especially the data type of the dependent variables, and the possible relationship between the dependent variable and independent variables.

- If the dependent variable is continuous, one can choose between simple linear regression and general linear regression.
  - If one wants to determine the linear relationship between one independent variable and one dependent variable, simple linear regression should be used.
  - If one needs to include multiple independent variables in the model, general linear regression can be used.
- If the outcome is discrete, for instance, the patient is alive or dead after the treatment, one can use logistic regression.

For a given data set, to determine the relationship between one independent variable and one dependent variable, one can draw a scatterplot, as shown in Figure 8.8.

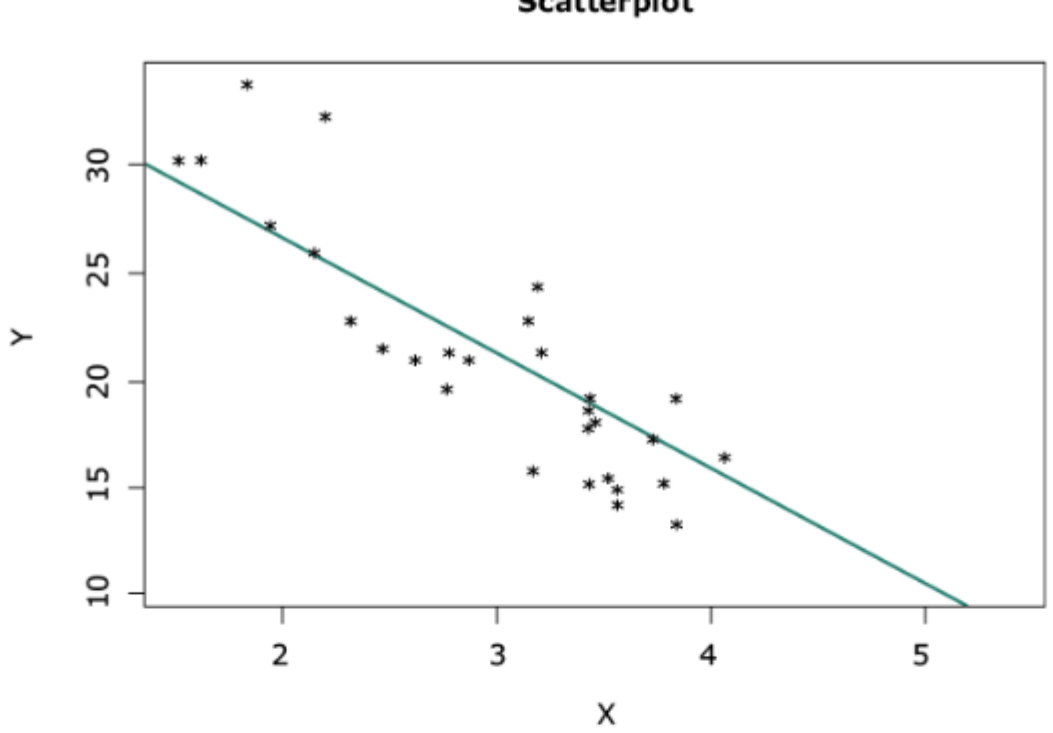


Figure 8.8. Scatterplot to show the relationship of two variables

▶ Expand/collapse Extended Description

Big Data Analytics

page 225

As mentioned in the first section of this chapter, one of the characteristics of big data is that there is a wide variety of data types available for analysis. Examples include EHRs; lab test results; digital images from CT, MRI, ultrasound, and x-rays; wearable sensor data; patient reported outcomes; genomic data; answers to questionnaires; and data collected by mobile health apps. One project of this type is the **All of Us Research Program** created under the Precision Medicine Initiative (Sankar and Parker 2017). The All of Us Research Program will create a research cohort of 1 million or more Americans who have volunteered to share their genetic data, biological samples, and diet and lifestyle information, all linked to their EHRs (Collins and Varmus 2015). The goals include:

- Advancing pharmacogenomics, the right drug for the right patient at the right dose
- Identifying new targets for treatment and prevention
- Testing whether mobile devices can encourage healthy behaviors
- Laying a scientific foundation for precision medicine for many diseases (NIH 2022)

To analyze such a massive and heterogenous data set, big data analytics skills need to be applied. First, all the raw data from different sources need to be extracted, transformed into a consistent format, and loaded into cloud-based management programs. In most cases, data analysts need to build their own programs to handle this task because it is new to everyone. After the transformation is done and the data are in the cloud, some software tools for big data analysis, such as Hadoop, MapReduce, and Hive, can be used to prepare subsets of data for further analysis with more specific purposes. In this last step, many regular data analysis tools and skills described earlier (such as data visualization,

descriptive statistics, inferential statistics, and statistical modeling) can be applied on those subsets of data.

There are three major challenges in big data analytics. The first one is the quality of data. In the All of Us Research Program, some data items may have high quality, such as the EHRs, the lab test results, and the genetic data, because they are handled by professionals with specific training based on predetermined standards. Some data, such as diet and lifestyle information, might be of poor quality because they are either reported by participants or collected via wearable sensors. Poor-quality data will not generate high-quality results.

The second challenge is to integrate highly heterogenous data together. For instance, what is the best way to integrate wearable sensor data with genomic data? This is still an open research question in big data analytics.

The third major challenge is the volume of the data. Let us take the genomic data as an example. If we ignore everything else and just focus on the genome sequence itself, the size of one genome is 3 gigabytes (GB). When it is multiplied by 1 million, the sequence data alone will be at the scale of petabytes (PB), which has never been handled by most data analysts in healthcare.

Besides these three major challenges, there are several other challenges in big data analytics, including security and privacy issues. In the All of Us Research Program, participants' EHRs, demographics, genetic data, and their lifestyle data are collected. When all these pieces of deidentified data are integrated together, it is possible to reidentify individual participants. The discussion of this topic is beyond the scope of this chapter.

Results Presentation

Presenting the results of a data analysis effectively can be a challenge. The data output from an analysis might or might not be clear or easy to understand. When numbers are displayed incorrectly, that can encourage false conclusions from the analysis results. Whether presented with words, in tabular form, or in a graph or picture, it is imperative the presentation be clear.

Tables

A table is an organized arrangement of data, usually in columns and rows. Many different types of quantitative data can be displayed in tables. For examples of tables see tables 8.1 through 8.4. Tables should be used with care because the inclusion of too much data in a table can increase rather than decrease confusion when presenting results.

If the decision is made to use a table, it is helpful to keep the following guidelines in mind:

- The table should be a logical unit that is self-explanatory and stands on its own.
- The source of the data in the table should be specified.
- Headings for rows and columns should be understandable.
- Blank cells should contain a zero or a dash.
- Formatting for headings and cell contents should be consistent so that the eye is not confused (Marc 2016, 540).

In some instances, a table can be the basis for a chart or graph to display the data. The data visualization techniques and guidelines can still be applied to the preparation of results presentation. The following rules can be useful for the data visualization at this stage.

1. Know your audience.
2. Identify the key message you want to deliver.
3. Use color effectively.
4. Adapt the diagram to the presentation medium (for example, a piece of printed paper, a large poster, computer screen, projector, and such).
5. Do not mislead the reader.
6. Avoid "chartjunk," or unnecessary details in the diagram.
7. Give message a higher priority than beauty of the diagram.
8. Use the right tool for creating the data visualization. (Rougier et al. 2014)

The steps described for data analysis are the typical steps data analysts will go through when they start a new project, especially in recent years, because it has become much easier to collect various types of healthcare data. On the other hand, one does not necessarily need to go through all the steps described earlier. In some cases, when the data are well organized, and the quality of data is high, data analysts might not need to perform much data preprocessing but may directly start EDA. In other cases, data analysts might need to spend a significant amount of time on data preprocessing if the data are extremely messy and the data quality varies widely. Beyond statistical modeling, data analysts might also use other data analytics techniques on the data sets, such as applying machine learning algorithms to the data and making predictions from the data, which is the topic covered in detail in the next chapter.

Case Study: National Health and Nutrition Examination Survey

The NHANES is a program designed to assess the health and nutritional status of adults and children in the United States. This program began in the early 1960s and a number of surveys have been conducted since then. The surveys include interviews and physical examinations. The interviews collect data such as demographics, socioeconomic status, dietary, and health-related issues. The physical examinations collect medical, dental, and physiological data, and laboratory test results. Each year, about 5,000 persons identified nationwide participate this survey. The collected data are deidentified and made publicly available. Therefore, students can access the data and conduct a wide variety of analyses. For instance, the NHANES data can be used to figure out the prevalence of major diseases and risk factors for diseases. The findings from the data analysis can be used to inform decision-making at the national level.

Case Study: MIMIC-III Data Set

The Medical Information Mart for Intensive Care (MIMIC-III) is a database comprising deidentified health-related data associated with more than 40,000 patients who stayed in critical care units of the Beth Israel Deaconess Medical Center between 2001 and 2012 (Johnson et al. 2016). Although the data set itself is freely available, because MIMIC-III contains

detailed information about real patients' care, it requires credentialing before access. To help data analysts to determine whether the MIMIC-III data set is useful for their studies before requesting full access to the database, a demo subset was created with data for 100 patients (Johnson et al. 2016). Data analysts can use this freely available demo set to perform missing data handling, data transformation, and data visualization so that they can have a better idea about the data set.

Chapter 8 Review Exercises

Indicate whether the following statements are true or false.

1. Data visualization methods can be applied to medical, healthcare, and geographic information systems and census and personal health data.
2. Salaries collected in ranges of 0–25,000, 26,000–30,000, and such, are continuous data.
3. Imputation procedures for a research project do not need to be documented in the resulting paper.
4. Replacing too much data by imputation can bias your study.
5. Data standards are needed to facilitate interoperability of diverse systems.
6. Structured data can come from webpage drop-down lists and check boxes.

Discussion

7. Describe the types of data, as well as the relationship between data and information.
8. Compare and contrast the methods of collecting structured and unstructured data. List one disadvantage of each type of data.
9. Design an appropriate data presentation for displaying the length of stay in the transfers table in the MIMIC-III demo data set. Justify the choice of your graph.
10. Explore the NHANES data set and propose a data science project, indicating the variables that might be useful for the project. There are many variables in this data set.
11. Explain what big data are and how they are used.

REFERENCES

AHIMA. 2017. ***Pocket Glossary of Health Information Management and Technology***, 5th ed. Chicago: AHIMA Press.

Ash, J. S., D. F. Sittig, E. M. Campbell, K. P. Guappone, and R. H. Dykstra. 2007. Some unintended consequences of clinical decision support systems. ***AMIA Annual Symposium Proceedings*** 2007:26–30.

Centers for Disease Control and Prevention (CDC). 2020. "***NHANES 2017-2018 laboratory data.***" https://wwwn.cdc.gov/nchs/nhanes/search/datapage.aspx?Component=Laboratory&CycleBeginYear=2017

Centers for Disease Control and Prevention (CDC). 2022. "***Mortality statistics.***" http://www.cdc.gov/nchs/nvss/deaths.htm.

Cokluk, O., and M. Kayri. 2011. The effects of methods of imputation for missing values on the validity and reliability of scales. ***Educational Sciences: Theory and Practice*** 11(1):303–309.

Collins, F. S., and H. Varmus. 2015. A new initiative on precision medicine. ***New England Journal of Medicine*** 372(9):793–795. doi:10.1056/NEJMp1500523.

Endicott, M. 2014. Transitioning from E codes in ICD-9-CM to V, W, X, Y codes in ICD-10-CM. ***Journal of AHIMA***. http://journal.ahima.org/2014/08/13/monthly-code-cracker-transitioning-from-e-codes-in-icd-9-cm-to-v-w-x-y-codes-in-icd-10-cm/ .

Fenton, S. H. 2006. "Structured or unstructured? Options for clinician data entry in the EHR." ***American Health Information Management Association***. http://bok.ahima.org/doc?oid=63115.

Gartner, Inc. 2013. "Big data definition." ***IT Glossary***. http://www.gartner.com/it-glossary/big-data/.

Green, S. B., and N. J. Salkind. 2016. ***Using SPSS for Windows and Macintosh***, 8th ed. Upper Saddle River, NJ: Pearson.

Hugine, A. L., S. A. Guerlain, and F. E. Turrentine. 2014. Visualizing surgical quality data with treemaps. ***Journal of Surgical Research*** 191(1):74–83. doi: 10.1016/j.jss.2014.03.046.

Jardine, A., N. Mullan, O. Gudes, J. Cosford, S. Moncrieff, G. West, X. Jianguo, G. Yun, and P. Somerford. 2014. Web-based geo-visualisation of spatial information to support evidence-based health policy: A case study of the development process of HealthTracks. ***Health Information Management Journal*** 43(2):7–16. doi:10.12826/18333575.2014.0004.

Johnson, A. E. W., T. J. Pollard, L. Shen, L. H. Lehman, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. A. Celi, and R. G. Mark. 2016. MIMIC-III, a freely accessible critical care database. ***Scientific Data*** 3:160035. https://physionet.org/content/mimiciii-demo/1.4/.

Johnson, S. B., S. Bakken, D. Dine, S. Hyun, E. Mendonça, F. Morrison, T. Bright, T. Van Vleck, J. Wrenn, and P. Stetson. 2008. An electronic health record based on structured narrative. ***Journal of the American Medical Informatics Association*** 15(1): 54–64. doi:10.1197/jamia.M2131.

Komenda, M., M. Víta, C. Vaitsis, D. Schwarz, A. Pokorná, N. Zary, and L. Dušek. 2015. Curriculum mapping with academic analytics in medical and healthcare education. ***PLoS ONE*** 10(12):1–18. doi:10.1371/journal.pone.0143748.

Marc, D. 2016. Data Visualization. Chapter 18 in ***Health Information Management Concepts Principles, and Practices***, 5th ed., edited by P. Oachs and A. Watters. Chicago: AHIMA.

Mishra, P., C. M. Pandey, U. Singh, A. Gupta, C. Sahu, and A. Keshri. 2019. Descriptive statistics and normality tests for statistical data. ***Annals of Cardiac Anaesthesia*** 22(1):67–72. https://doi.org/10.4103/aca.ACA_157_18.

National Institutes of Health (NIH). (2022). "The Precision Medicine Initiative." https://www.nih.gov/sites/default/files/research-training/initiatives/pmi/pmi-infographic.pdf.

National Library of Medicine (NLM). 2022. "Value Set Authority Center." https://vsac.nlm.nih.gov/.

Rosenbloom, S. T., H. Denny, N. Xu, N. Lorenzi, W. W. Stead, and K. B. Johnson. 2011. Data from clinical notes: A perspective on the tension between structure and flexible

documentation. ***Journal of the American Medical Informatics Association*** 18(2):181–186. doi: 10.1136/jamia.2010.007237.

Rougier, N. P., M. Droettboom, and P. E. Bourne. 2014. Ten simple rules for better figures. ***PLOS Computational Biology*** 10(9):e1003833. doi: 10.1371/journal.pcbi.1003833.

Safran, C., M. Bloomrosen, W. E. Hammond, S. Labkoff, S. Markel-Fox, P. C. Tang, and D. E. Detmer. 2007. Toward a national framework for the secondary use of health data: An American Medical Informatics Association white paper. ***Journal of the American Medical Informatics Association*** 14(1):1–9. doi:10.1197/jamia.M2273.

Sankar, P. L., and L. S. Parker. 2017. The Precision Medicine Initiative's All of Us Research Program: An agenda for research on its ethical, legal, and social issues. ***Genetics in Medicine*** 19(7):743–750. doi:10.1038/gim.2016.183.

Shaw, P. L., and D. Carter. 2015. ***Quality and Performance Improvement in Healthcare: Theory, Practice, and Management***, 6th ed. Chicago: AHIMA.

Shneiderman, B., C. Plaisant, and B. W. Hesse. 2013. Improving healthcare with interactive visualization. ***Computer*** 46(5):58–66. doi: 10.1109/MC.2013.38.

US Census Bureau. 2020. "***2020 Census illuminates racial and ethnic composition of the country.***" https://www.census.gov/library/stories/2021/08/improved-race-ethnicity-measures-reveal-united-states-population-much-more-multiracial.html .

van Mulligen, E. M., H. Stam, and A. M. van Ginneken. 1998. Clinical data entry. ***Proceedings / AMIA Annual Symposium. AMIA Symposium***, 81–85.

ADDITIONAL RESOURCES

Centers for Disease Control and Prevention (CDC). 2013. ***International Classification of Diseases, 10th rev, Clinical Modification (ICD-10-CM)***. http://www.cdc.gov/nchs/icd/icd10cm.htm.

Centers for Medical and Medicaid Services. 2016. ***ICD-10-PCS***. https://www.cms.gov/Medicare/Coding/ICD10/downloads/pcs_final_report2012.pdf.

Department of Health and Human Services (HHS). 2012. "Administrative simplification: Adoption of a standard for a unique health plan identifier; Addition to the National Provider Identifier requirements; and a change to the compliance date for the International Classification of Diseases, 10th ed. Final Rule." ***Federal Register.*** https://www.federalregister.gov/documents/2012/09/05/2012-21238/administrative-simplification-adoption-of-a-standard-for-a-unique-health-plan-identifier-addition-to .

Department of Health and Human Services (HHS). 2015. "Health IT Standards Committee." https://www.healthit.gov/FACAS/health-it-standards-committee.

Health Story Project. 2013. http://www.healthstory.com/.

Ramanathan, T., C. Schmidt, A. Menon, D. Pepin, M. Penn. 2015. "Issue brief: Federal public health laws supporting data use and sharing." https://www.cdc.gov/phlp/docs/datasharing-laws.pdf.

Chapter

9

Data Analytics and Machine Learning

By Yanshan Wang, PhD, FAMIA, and Sonish Sivarajkumar

Learning Objectives

- Understand data analytics and its subtypes
- Use data analytics process to get better insights from healthcare data
- Apply different steps in data analytics process
- Interpret the use of machine learning in data analytics pipeline
- Evaluate the performance of a machine learning model

KEY TERMS

Accuracy
Area under the curve (AUC)
Classification
Clustering
Confusion matrix
Data analytics
Data lake
Data preprocessing
Data warehouse
Density estimation
Descriptive analytics
Diagnostic analytics
F1-score
False negative (FN)
False positive (FP)
Feature Engineering
K-means clustering

K-nearest neighbor
Linear regression model
Logistic regression
Machine learning
Mean absolute error
Mean squared error (MSE)
Naïve Bayes
Named entity recognition
Neural networks
Outlier detection
Positive Predictive Value (PPV)
Precision
Predictive analytics
Prescriptive analytics
Principal component analysis (PCA)
R^2 (R-squared)
Recall
Receiver operating characteristic (ROC)
Regression
Reinforcement learning
Sensitivity
Specificity
Support vector machine (SVM)
Supervised learning
True negative (TN)
True positive (TP)
Unsupervised learning

Introduction to Data Analytics

Data analytics refers to the use of data, information technology (IT), and statistical analysis to gain better insights into a problem, and thus make better fact-based decisions. IBM defines data analytics as "the systematic use of data and related business insights developed through applied analytical disciplines to drive fact-based decision making for planning, management, measurement, and learning." In other words, data analytics can be defined as the process of identifying and communicating important patterns in data.

Recent years have witnessed health care evolving with the adoption of technology. The data collected digitally from hospitals and healthcare providers has shown the potential to improve patient care and facilitate better decision-making processes. Together with patient care, the use of data analytics also allows optimization of the operational processes of businesses.

Types of Analytics

There are four different types of analytics: descriptive analytics, diagnostic analytics, predictive analytics, and prescriptive analytics, as described next.

Descriptive analytics is mainly used for reporting by analyzing historical data. This type of analytics usually helps answer questions, such as "What is happening?" or "What has happened?" It is useful in understanding more in-depth insights into specific queries with the data collected from the past. Descriptive analytics can be used to get an

aggregate-level understanding of an environment in a healthcare organization, like how many patients were admitted to the hospital, and in which department they were admitted.

Diagnostic analytics is used for diagnosing an event that has already occurred and finding the reason for the event. In other words, this type of analytics can answer the question, "Why did this happen?" Diagnostic analytics will be able to find the cause of an issue based on the collected descriptive data; for instance, what is the cause of death of many patients in an intensive care unit (ICU)?

Predictive analytics uses statistical models and forecasting techniques to identify trends, correlation, and causation. Predictive analytics estimates the likelihood of a future event by using predictive models. It can be used any time you need to know something about the future; for example, "What will happen?" Predictive analytics can be used to identify the patients who will likely be readmitted to hospitals in the next 30 days by analyzing their medical history.

Prescriptive analytics helps to identify the course of action to achieve something by leveraging optimization and simulation algorithms to advise on possible outcomes. Prescriptive analytics basically answers the question, "What should we do?" For example, what is a more cost-effective way to manage patients at high risk for readmission: at the hospital or at the patient's home? Predictive analytics recommends one or more possible courses of action. Due to the complexity of prescriptive analytics, it has not been widely adopted in routine care. However, prescriptive analytics has become a hot research topic in health informatics (Davenport 2013).

Steps of Data Analytical Process

The data analytics process starts with collecting the data and moves to performance evaluation and visualization of the system trained on these data. Figure 9.1 shows the flowchart of the data analytical process. Each step will be described in detail in the following sections.

Data Collection

There are numerous sources in the healthcare and biomedical domain that generate different types of data, such as hospitals, healthcare providers, pharmaceutical companies, and so forth. Those data could be structured (for example, demographics, diagnosis codes, medications, and lab tests) or unstructured (for example, clinical notes and radiology reports). Collecting these structured and unstructured data is the basic step of the data analytics pipeline. Data lakes and data warehouses are commonly used to store large amounts of data. A **data lake** is defined as "a methodology enabled by a massive data repository based on low-cost technologies that improves the capture, refinement, archival, and exploration of raw data within an enterprise" (Fang 2015). A data lake is a large collection of unstructured or structured raw data. This can be thought of as a central database for storing data of various types and veracities, from multiple sources. The stored data could be semistructured, unstructured, raw, or unprocessed. A **data warehouse** is a repository where the data, which have been processed for a specific purpose, can be stored. Thus it is a collection of structured and processed data (Fang 2015).

Three major issues can emerge during data collection. The first is missing data. There are many reasons for missing data that vary substantially across studies and organizations (Haneuse et al. 2021). For example, if a patient might have dropped out, failed to follow up, or was unrecorded for measurement during an office visit. The second issue is data imbalance. Some data classes or categories could have a relatively larger number of samples, which could be disadvantageous to other classes, leaving some individuals

underrepresented in the model. The third issue is data bias. For example, many human research studies are heavily biased toward participants with a Western, educated, industrialized, rich, democratic (WEIRD) profile, and are thus not representative of the human population as a whole (Norori et al. 2021).

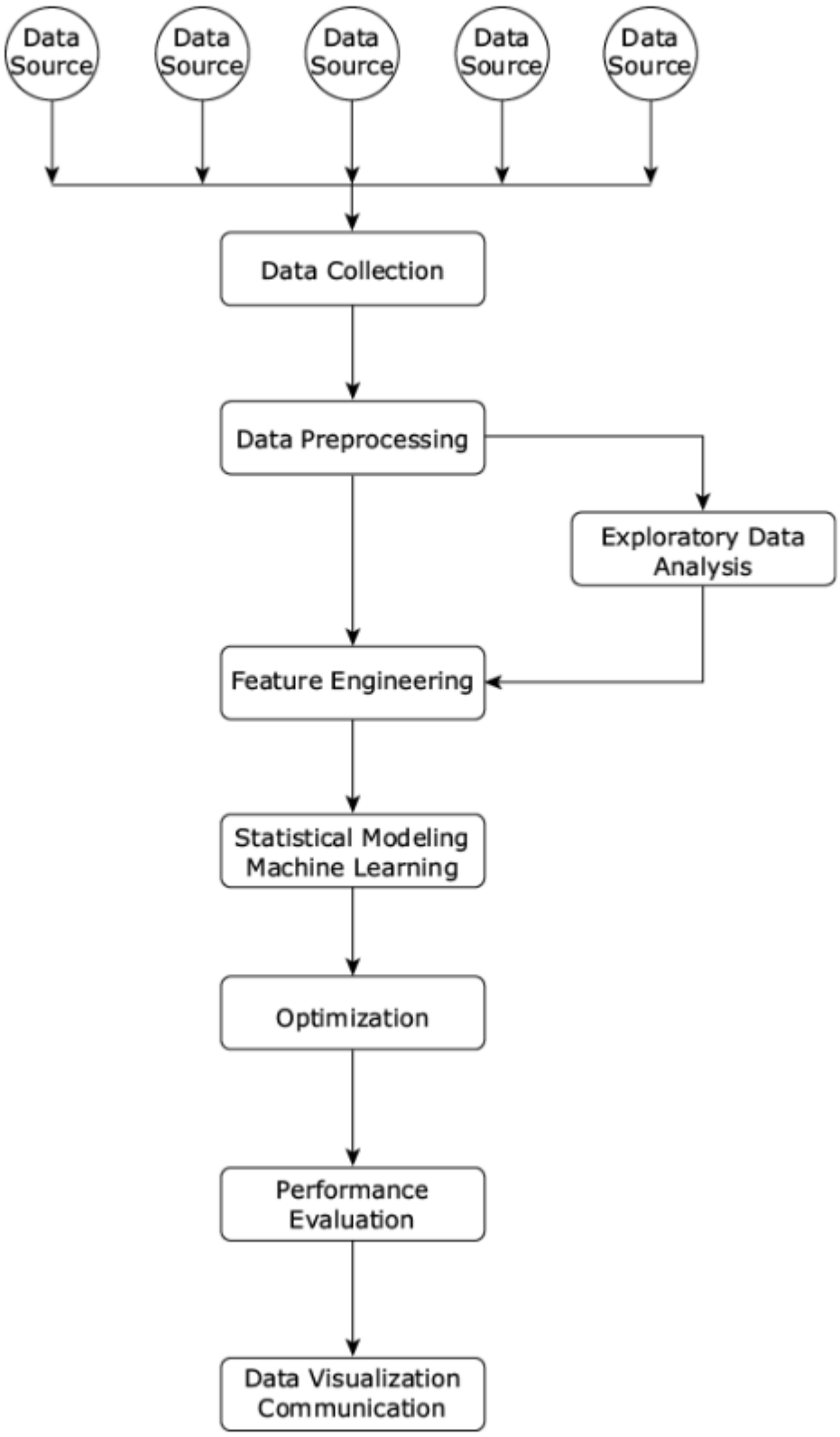


Figure 9.1. Steps of the data analytical process

▶ Expand/collapse Extended Description

Preprocessing

Data preprocessing refers to the transformation of incomplete, inconsistent, and imbalanced data into a usable form. This process includes a series of techniques that manipulate raw and noisy data.

Data Cleaning

The raw data in the database might include erroneous or irrelevant components that need to be cleaned. Data cleaning is the process of preparing raw data for analysis by removing these data errors. The data errors might be detected as an inconsistent or missing reading in a series of measurements. For example, inconsistent measurements on a machine-generated data could be the result of a damaged sensor. Data cleaning includes removing unnecessary data or outliers, removing duplicates, and so on (García et al. 2015). The data can be duplicated due to human error and improper data management. The demographic record of a readmitted patient might be entered twice in the hospital database, for example, resulting in a duplicate record. A domain expert can determine whether the data are necessary or not. Statistical measurements like standard deviation can also be used to identify outliers.

Missing Data Handling and Imputation

The raw data may have missing values, a common problem in almost all data analytics projects. The missing data can be a source of data imbalance and data bias in a data set. Therefore, it is critical to handle the missing data. Deletion, mean substitution, and regression imputation are commonly used to handle missing values. For deletion, one can simply delete the data points that have the missing values. However, this could result in the loss of data, which is critical when the data set is small. In mean substitution, the mean values of a variable are used in place of the missing values in that variable. This method assumes that the mean is a realistic estimate for a randomly selected observation from a normal distribution. However, mean substitution may result in an inconsistent bias when missing values are not strictly random, especially when there is a large disparity in the number of missing values for the different variables. Regression imputation is a comprehensive technique that replaces the missing values with a probable value, estimated by statistical regression techniques using existing values (García et al. 2015).

Data Normalization

Features are the variables or attributes in a data set. A data set might have features like age, height, weight, and so on. Different features in a data set might have different scales. For instance, height could be expressed in feet, and weight could be expressed in kilograms. Let's assume that a patient has a height of 6 feet and a weight of 60 kilograms. Computing systems understand these as numerals 6 and 60, which means that the attribute "weight" may have a higher influence on the statistical model. To avoid this, we usually use statistical data transformation or data normalization methods to transform features to a similar scale (for instance, 0–1), thus reducing the data inconsistency and bias between the attributes.

Feature Engineering

The performance of a predictive model is influenced by the model, the available data, and the features. **Feature engineering** is the process of selecting, manipulating, and transforming raw data into features that can be used in the next phases of the data analytics process (Zheng and Casari 2018). These features will directly influence the performance of the predictive models. The better the relevant features that we prepare and select, the better the results we can achieve. Machine learning methods are sensitive to

the dimensionality of the data we use. The question we are addressing here is this: Is there a lower dimensional representation of the data that effectively captures its characteristics?

Consider the mortality rate prediction of neighborhood-based morphological and clinical data for the residents of the neighborhood. The numerical features like age, height, weight, and average blood pressure can be used to measure the perceived survival rate of the residents. You could think of this apparent survival rate as the result of a kind of feature engineering.

For a feature to be useful, it must have a relationship to the target that your model is able to learn. Linear models, for instance, are only able to learn linear relationships. Therefore, when using a linear model, the goal of feature engineering is to transform the features to make their relationship to the target linear.

The key idea here is that a transformation you apply to a feature becomes in essence a part of the model itself. For example, if you are trying to predict the survival rate of an individual from his body mass index (BMI), fitting a linear model directly to the BMI could give poor results if the relationship is not linear.

Clustering

The purpose of clustering is to learn some properties of the data, thus exploring their underlying features. **Clustering** groups the data together based on how similar the points are to each other. A clustering algorithm can categorize the patients with similar morphological features. Studies have proved that data clustering can be used as a way of extracting new features and can improve the quality of machine learning models (Goodfellow et al. 2016).

page 239

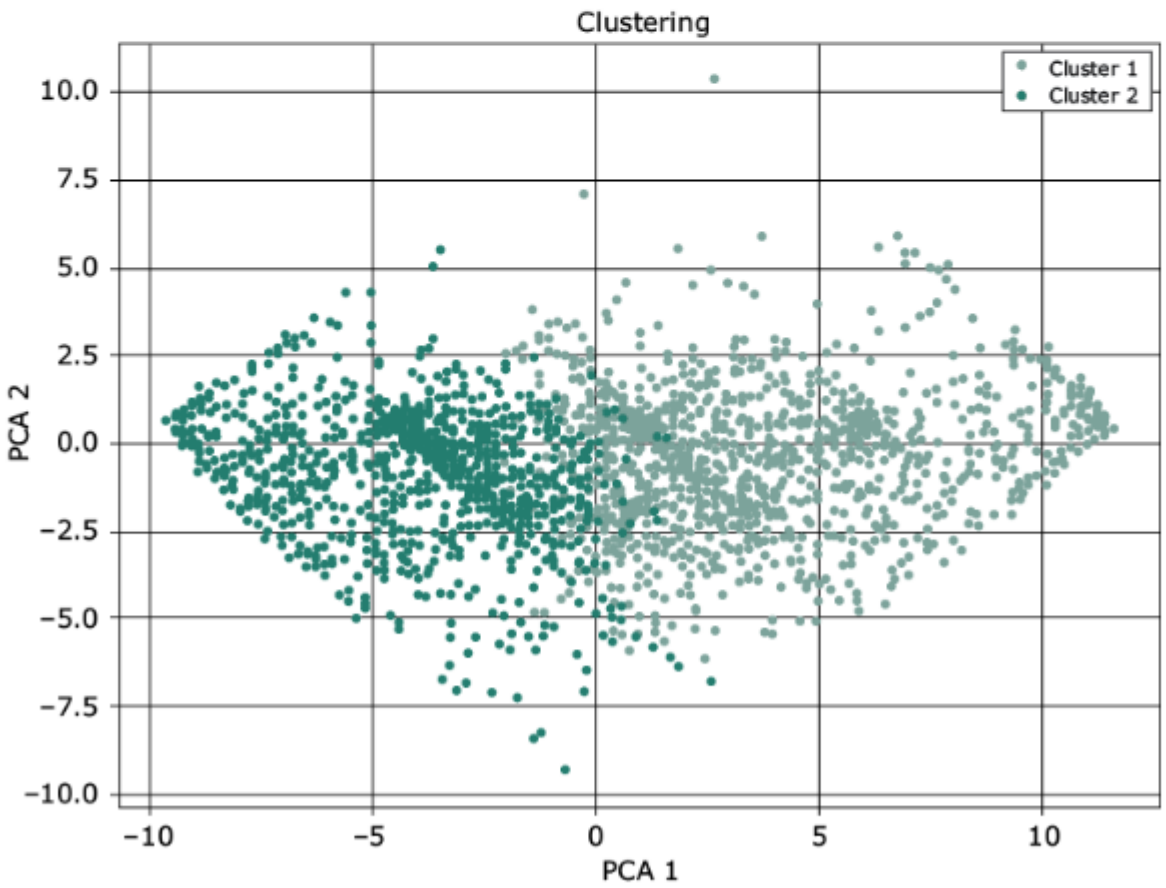


Figure 9.2. Clustering

▶ Expand/collapse Extended Description

In clustering, the data are distributed into different groups such that data points similar to each other are in the same group (see figure 9.2). In the context of feature engineering, we could attempt to discover groups of patients representing a disease condition, for instance, or geographic areas that share similar social determinants of health like economic status. Different clustering algorithms include K-means clustering, probabilistic (soft) clustering, and hierarchical (agglomerative) clustering. Details of clustering are described in later sections.

Principal Component Analysis

Principal component analysis (PCA) is another commonly used feature extraction method (Goodfellow et al. 2016). PCA projects the high-dimensional input data into lower-dimensional data by combining features. The data set thus gets smaller, with fewer important features. Just like clustering is a grouping of the data set based on distance or proximity, you could think of PCA as a partitioning of the variation in the data. It combines highly correlated variables together to form a smaller data set. In other words, PCA extracts the principal components of a data set to represent the most information in the original data.

Check Your Understanding 9.1

Match the terms with the appropriate descriptions

a. Descriptive analytics
b. Diagnostic analytics
c. Predictive analytics
d. Prescriptive analytics

1. _____ How should we take action?
2. _____ Where are we going and when?
3. _____ Why have we seen the past results?
4. _____ What have we done in the past?

Indicate whether each of the following is true or false.

5. The purpose of data preprocessing is to clean, correct, remove, and reduce data.
6. Clustering can be considered a data cleaning process.
7. Data normalization brings all the features of a data set to a common scale.
8. PCA can be used to group together patients with similar phenotypes.

Answer the following questions.

9. Which of the following processes involves grouping together similar data points that are positioned or occurring closely together?
   a. Data collection
   b. Data analytics
   c. Clustering
   d. Principal component analysis (PCA)
10. Which technique is used for reducing the dimensionality of large data sets by splitting a large collection of variables into a smaller one that still retains most of the information in the large set?
    a. Data collection
    b. Data analytics
    c. Clustering
    d. Principal component analysis (PCA)

11. Data sets with multiple features having values on different scales could lead to a poor data analytics system as it gets biased toward features with values on a larger scale. Which of the following processing steps involves bringing all the features to a common scale?
    a. Data collection
    b. Data normalization
    c. Data cleaning
    d. Data warehousing

Model Training: Machine Learning

Machine learning (ML) is a subset of Artificial Intelligence(AI) technology in which algorithms scan vast data sets to detect patterns, learn from them, and do daily jobs that humans do effortlessly. The broad availability of powerful technology and cloud computing has increased the acceptance of machine learning in various areas of human life, including the healthcare industry.

Due to the vast amount of data collected for patients and providers, machine learning algorithms in healthcare have a great potential to assist or automate routine tasks in healthcare domain. Consequently, the healthcare industry has already adopted many machine learning-based applications, including advanced patient management and clinical decision support systems.

Traditional computer systems produce output from a set of rules specified by the user. **Machine learning** is a technology whereby computers are trained to make decisions or predictions based on given data. Computers learn patterns in the data to make informed decisions in the future, as demonstrated in figure 9.3.

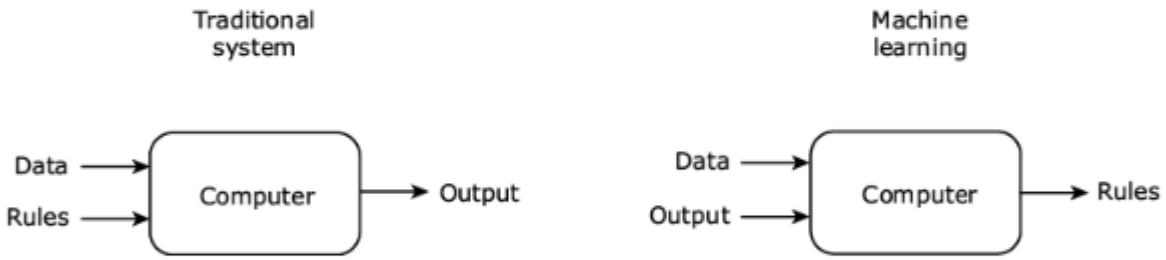


Figure 9.3. Traditional system versus machine learning

▶ Expand/collapse Extended Description

Types of Learning

Based on the type of the data and learning methodology, machine learning is classified into three separate categories:

- Supervised learning
- Unsupervised learning
- Reinforcement learning

Supervised Learning

Supervised learning takes the data as pairs (input, output) and learns a mapping function that could calculate the output, sometimes also called the response or target, from the input. In supervised learning, we train the model using labeled data that have been tagged with one or more truth labels by humans. Once the supervised model is trained on the labeled data, it returns the correct outcome when we pass the new unlabeled input data.

For instance, we train a model giving BMI as the input and the health status as the label. Once the model has learned the patterns, the model can predict the health status of a patient given his or her BMI (Bas¸tanlar and Özuysal 2014; Goodfellow et al. 2016).

Supervised learning can be further classified into two broad categories of algorithms. **Regression** is used when the target variable is numerical, such as income, age, and so forth. A regression model seeks the mathematical relationship between two or more numerical variables. For example, linear regression can be used to predict the total deaths in a locality given the number of patients infected with COVID-19. The most common regression model is a **linear regression model,** which fits the training as close to the line as possible. The line is then used to predict the target for unseen data. See figure 9.4 for an example of a linear regression model.

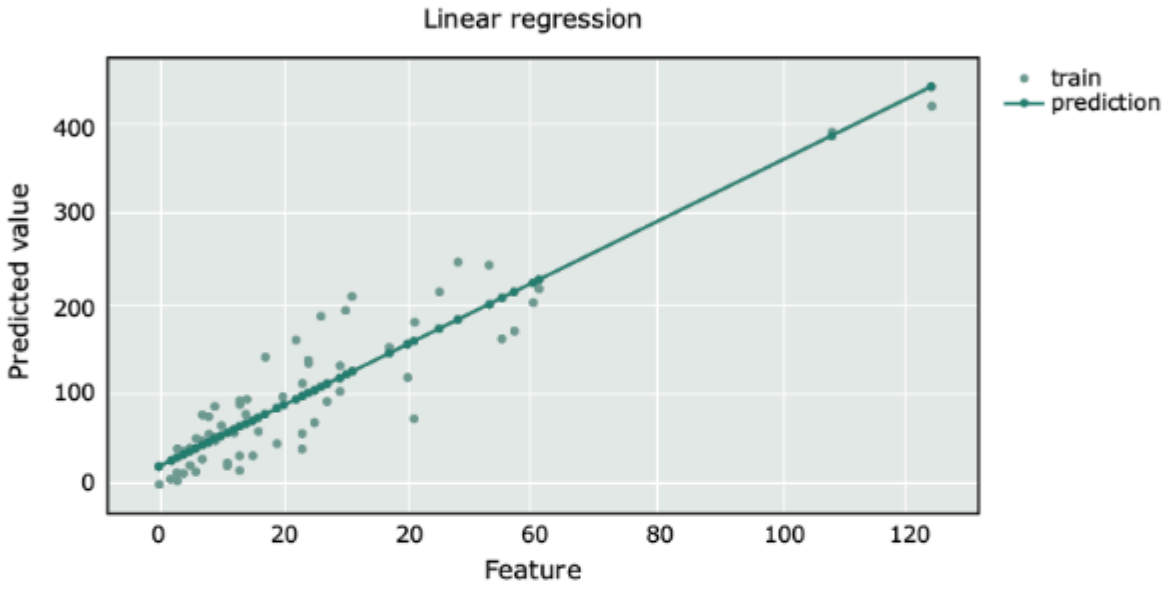


Figure 9.4. Linear regression

▶ Expand/collapse Extended Description

Classification is used when the target is nominal or categorical, such as cancer versus no cancer or healthy versus diseased. Unlike regression where predictions are modelled as continuous valued functions, classification predicts categorical labels. Another example of classification might include normal, mild, moderate, and severe conditions.

The following are five classification algorithms:

- **Support vector machines (SVMs)** separate the attribute spaces into hyperplanes and determine the best coefficients for the separation of the data. These can be used for classification and regression.
- **Logistic regression** is used for classification, usually with a binary outcome such as alive (1) or dead (0). The outcome is reported as odds ratios, so the higher the odds the more likely a predictor (independent variable) will predict an outcome (dependent variable).
- **Neural networks** are organized like human neurons, but this approach is more empirical and more difficult to explain and confirm.
- **Naïve Bayes** uses the Bayes theorem to predict future probability, based on known prevalence. This algorithm is considered naïve because it assumes each variable is independent, which is not always the case.

- **K-nearest neighbor** works by looking for nearest neighbors using several methods, including Euclidean distance.

Supervised learning has been employed in a wide range of applications. Early detection of diseases using classification algorithms has helped to reduce the number of readmissions in hospitals and clinics. Pattern or future case prediction is another common application of supervised learning. We apply the rule or the pattern learned from the training data set to forecast the output for future inputs. This could be used, for example, to predict the demand for a drug in a geographic location based on the number of people in different age groups and their health status. Supervised learning is widely used in knowledge extraction as well. This includes text classification or text extraction from large clinical data sets. **Named entity recognition** is another common application in which we identify different entities from a sentence, like names, locations, and addresses. A simple example, NER may recognize the word 'Alzheimers' in a document and apply the label 'disease' to it. Another application of supervised learning is in **outlier detection**, where we detect anomalies in a data set, as in spam detection in emails. Patient management and alerting systems use outlier detection methods to alert in case of abnormal patient state.

Unsupervised Learning

Unsupervised learning refers to the use of algorithms to identify patterns from unlabeled data. The unsupervised learning model groups input data into clusters based on their common characteristics. Patient segmentation is a simple use case of unsupervised learning in healthcare. In this application, patients can be categorized based on their age groups or disease.

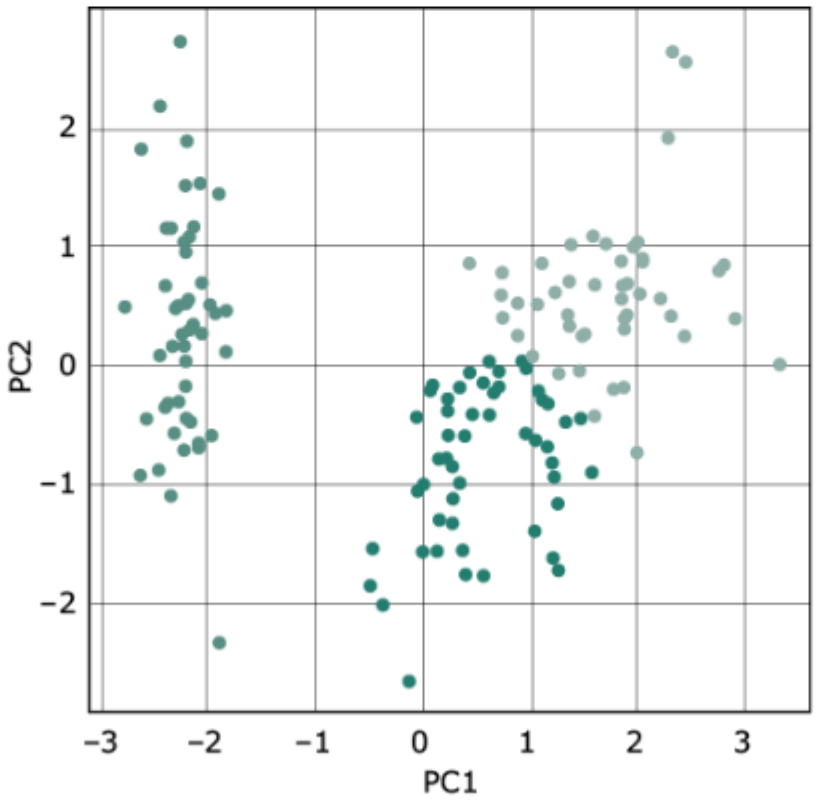


Figure 9.5. K-means clustering

▶ Expand/collapse Extended Description

There are two common forms of unsupervised learning. The first form is clustering, where we group together similar examples, for example, patient cases with common disease conditions. The second unsupervised learning method is **density estimation**. In density estimation, we probabilistically model the population of samples in a continuous density field. We estimate the probability density of a point in the two-dimensional space (for example, bin selection to fit a histogram on a probability distribution curve).

K-means clustering is an unsupervised learning algorithm. In this centroid-based or distance-based technique, the distances between points are calculated from a randomly assigned centroid. Each cluster in K-means is paired with a centroid. The centroids are moved based on the least within-cluster sum of distances. An example is displayed in figure 9.5.

Reinforcement Learning

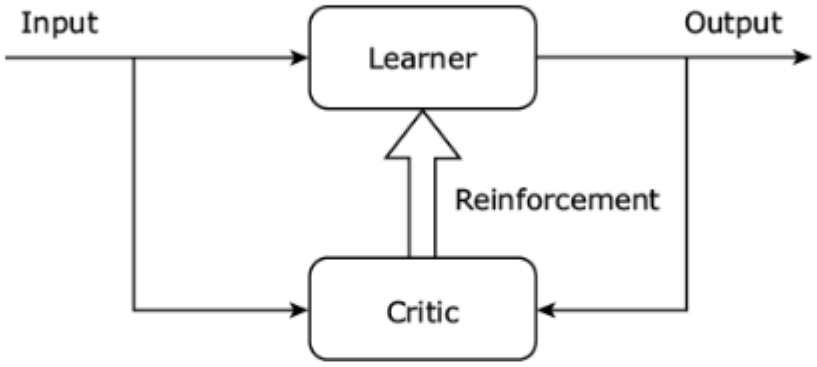


Figure 9.6. Reinforcement learning

▶ Expand/collapse Extended Description

There are other forms of learning apart from supervised, unsupervised, and reinforcement learning. These include advanced AI techniques like active learning, transfer learning, deep learning, and more.

Case Study: Disease Prediction

Accurate analysis and detection of diseases are important to identify patients who are at risk of a disease. This also facilitates preventive care and better treatment of the disease condition. Data analytics proves to be effective in assisting with decision-making and prediction from the large quantity of data produced by the healthcare industry. Developing a medical diagnosis system based on advanced data analytics and machine learning algorithms can help in predictive diagnosis of the illness. This case study describes how machine learning algorithms are used to identify patients with rare diseases.

Rare diseases often go undiagnosed, as they are too infrequent and found in an extremely low number of people. Acute hepatic porphyria (AHP) is one such rare disease, which causes life-threatening attacks in patients. This is a genetic disease, with 1 in 100,000 people with the gene mutation exhibiting signs and symptoms of the disorder. Hence, it is difficult for doctors to predict who will develop AHP in a given population.

Researchers from Oregon Health and Science University trained a machine learning model on 205,571 records in their medical system. They

identified 22 patients with symptoms of AHP. Those patients could then be clinically tested to find any occurrence of AHP.

This case study clearly describes how machine learning approaches can be used to detect an affected patient in a large population. Because testing for rare diseases is often expensive, predictive diagnosis helps to filter potential patients and provide preventive treatments to them.

(Source: Hersh et al. 2022)

Check Your Understanding 9.2

Match the terms with the appropriate descriptions

a. Supervised learning
b. Unsupervised learning
c. Reinforcement learning

1. _____ is a type of machine learning where the system is trained with unlabelled data.
2. Predicting the amount of rainfall in a region based on various cues is a _____ problem.

page 245

3. _____ is a type of machine learning that learns by interacting with its environment. The agent receives rewards by performing correctly and penalties for performing incorrectly.

Indicate whether the following statements are true or false.

4. Unsupervised learning requires target labels in the data set.
5. K-means clustering can be used for classifying case and control groups.
6. Text classification is a supervised learning task, given the labels.
7. Regression can be used for forecasting drug demand in a geographic location.

Indicate which type of method (r = regression; c = classification, l = clustering) the following algorithms fall under

8. _____ Logistic regression
9. _____ K-means clustering
10. _____ Linear regression

Answer the following questions.

11. Which of the following statements is true about machine learning?
    a. Machine learning extracts patterns in the raw data.
    b. Machine learning allows systems to learn from experience, without being explicitly programmed.
    c. Machine learning reduces human intervention on certain tasks.
    d. All of the above.
12. What are the two types of supervised learning?
    a. Classification
    b. Clustering
    c. Regression
    d. Progression
13. Which algorithms can be used both as an unsupervised learning algorithm and a feature reduction method?
    a. Regression
    b. Classification
    c. Clustering

d. Principal component analysis (PCA)

Performance Evaluation

Once a machine learning model is trained, the next step is the evaluation of the trained model. To evaluate the performance of the model, we divide the input data into training and testing data. This splitting of the data set is illustrated in figure 9.7. The training data contain samples of the input data, which are initially used to train the machine learning model, whereas the testing data are used to evaluate the model. The result of the performance evaluation of the machine learning model could be an evaluation score or a confusion matrix, which is explained later in this chapter. Sometimes, during training, validation data are also used to tune many parameters of the machine learning model, but this is not common in healthcare data analytics (Jiao and Du 2016).

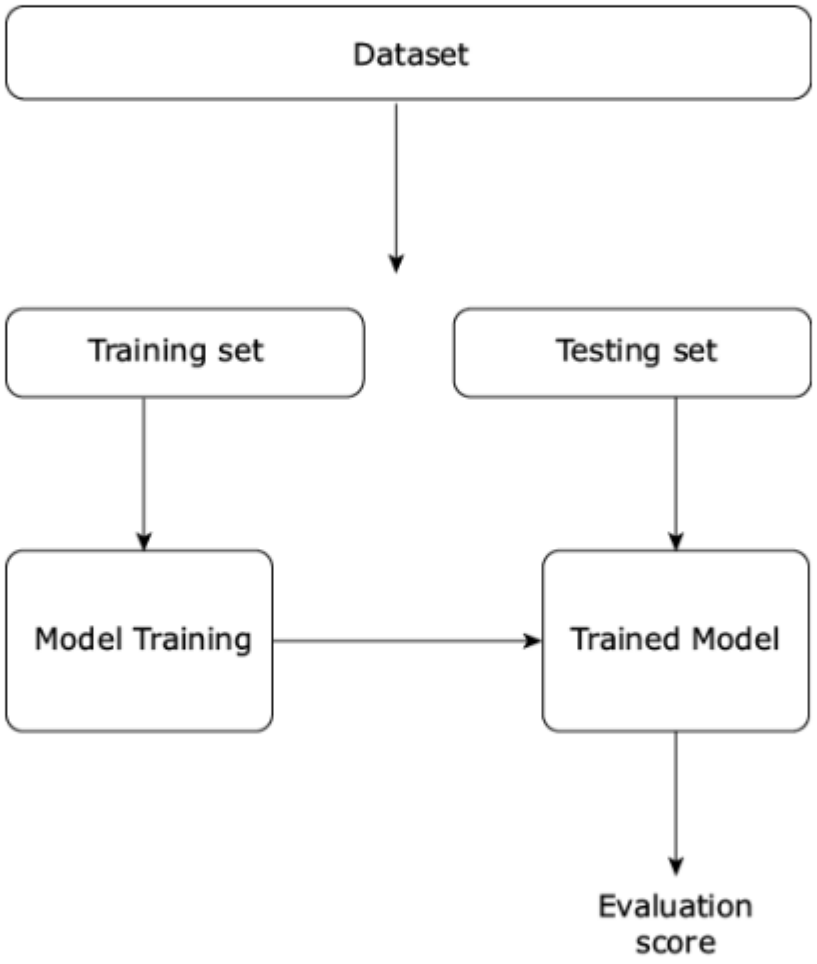


Figure 9.7. Data set splitting

▶ Expand/collapse Extended Description

Regression Model Evaluation

R^2 (R-squared) is a common measure to determine how close the fit to the observed values is. R^2 is the ratio of the sum of squares regression ($SS_{Regression}$) and the sum of squares total (SS_{Total}). Sum of squares regression ($SS_{Regression}$) represents the total variation of all the predicted values found on the regression line or plane from the mean value of all the values of response variables. R^2 determines how close the data fit the regression line with a range of 0 to 1. The closer the value is to 1, the better the fit.

$$R^2 = 1 - \frac{SS_{Regression}}{SS_{Total}} = 1 - \frac{\sum_i (y_i - y_i)}{\sum_i (y_i - \bar{y})^2}$$

Here, $\hat{y}$ is the predicted value, and y is actual value of the variable. represents the mean value of the variable.

Mean squared error (MSE) can be used to represent the cost associated with the prediction. This measures difference between the true value and the predicted value. When the linear regression model is trained using training data, the model with the lowest MSE value is selected as the best model. MSE is always positive. A perfect regression model will have zero MSE. A value closer to zero will represent better quality of the regression model.

$$MSE = \frac{1}{N}\sum_{i=1}^{N} \left(y_i - \hat{y}_i\right)^2$$

Mean absolute error (MAE) is similar to MSE. However, MAE takes the absolute value of the sum of errors, rather than its squared sum, as in MSE. It is a more direct representation of the regression error. Whereas MAE handles all errors equally, MSE squares high prediction errors and applies a bigger penalty.

$$MAE = \frac{1}{N}\sum_{i=1}^{N} \left|y_i - \hat{y}_i\right|$$

Classification Model Evaluation

Confusion Matrix

A classification model categorizes an input correctly or incorrectly. Every instance will have a ground-truth class (also called the gold standard class) and a predicted class (class predicted by the classification model). The ground truth is the true label of the data, whereas the predicted class is the output given by the machine learning system. These can be either positive or negative. There are four possibilities for every instance: true positive, true negative, false positive, and false negative.

A **true positive** (TP) is a case where the ground truth and the predicted label are of the positive class. When the model correctly classifies the presence of a disease, this is a true positive. A **true negative** (TN) is a case where the ground truth and the predicted label are negative, such as when the model correctly classifies the absence of disease. A **false positive** (FP) is a case when a model predicted a positive class, despite the fact that it was a negative class. This is the case when the classification model wrongly predicts a disease in a healthy patient. A **false negative** (FN) occurs when the model is a negative class, despite the fact that the ground truth is a positive class. For instance, it is a false negative if a test cannot diagnose a disease in a patient, which can be detrimental.

The **confusion matrix** is a table that visualizes the performance of classification models by showing the proportion of TPs, TNs, FPs, and FNs, as shown in figure 9.8. Evaluation metrics like precision, recall, accuracy, and F1-score can be easily obtained from a confusion matrix.

Accuracy

A general notion is that **accuracy** of a model can be used to evaluate the performance of a machine learning model, but accuracy can be misleading in many cases. For instance, in fraud detection or disease classification, a 90 percent accuracy may seem positive, but it might not consider the proportion of false positives or false negatives.

Accuracy can be formulated as follows:

$$Accuracy = \frac{TruePositive+TrueNegative}{TruePositive+TrueNegative+FalsePositive+FalseNegative}$$

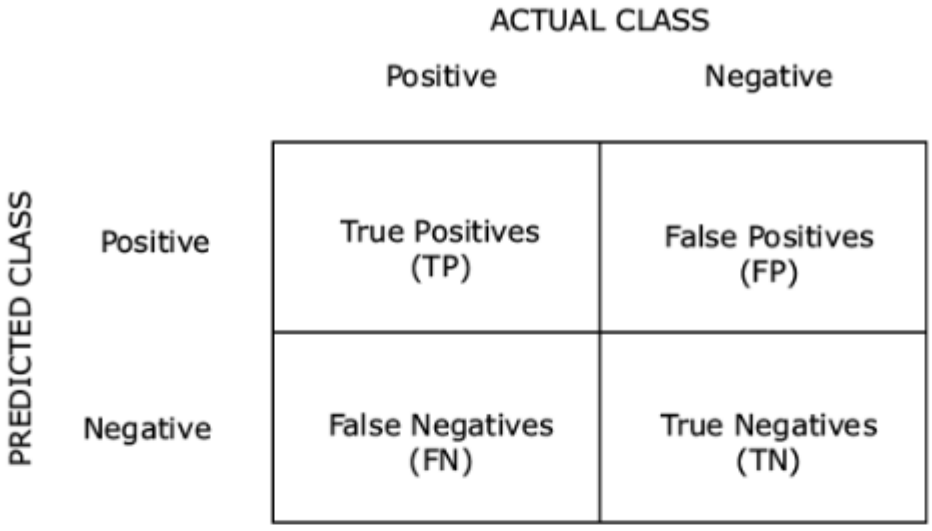


Figure 9.8. Confusion matrix

▶ Expand/collapse Extended Description

Let us consider a disease classification system being evaluated on 100 people, out of whom 98 are healthy and 2 are infected with the disease. The system might classify all 100 people as healthy, which gives an accuracy of 98 percent. Classifying the two infected patients, though, is the key task of the system. Hence, high accuracy alone does not mean a good system.

Precision

Precision or **positive predictive value (PPV)** is another evaluation metric for classification models. Mathematically, it can be written as:

$$Precision = \frac{TruePositive}{TruePositive+FalsePositive}$$

We use precision if the number of false positives is crucial. Consider a situation where a drug or therapy is administered to patients. It is essential to avoid usage of the drug by people who do not have that disease, meaning we need to avoid predicting a positive (predicting a disease) for a negative case (no disease). Because we want to avoid FP errors, we must have very high precision. Precision is the evaluation measure to be used to analyse the performance of the classification model in this scenario.

Recall

Recall is another evaluation measure for classification models. Recall is also known as **sensitivity** or **true positive rate** and is defined as follows:

$$Recall = \frac{TruePositive}{TruePositive+FalseNegative}$$

Recall is used when false negatives are critical to the application. It could be the best evaluation measure to analyze the performance of a classification model when false negatives are to be avoided. For instance, we need to avoid the situation when a patient is classified into a no-disease category in a disease detection model.

Specificity

Specificity is another common performance evaluation measure used for classification models. It is formulated as:

$$Specificity = \frac{TrueNegative}{TrueNegative+FalsePositive}$$

Specificity or true negative rate gives the performance of the machine learning system in predicting the negative class. In other words, when the value of specificity is high, it means that the system is good at predicting actual negative instances in the data set. In the application of detecting spam emails, high specificity means the system is good at correctly predicting spam emails correctly.

F1-Score

F1-score combines both precision and recall. This could be done by taking the mean of precision and recall. The harmonic mean is calculated as the number of values (N) divided by the sum of the reciprocal of the values.

$$HarmonicMean = \frac{N}{1/x1+1/x2+...+1/xN}$$

If there are just two values ($x1$ and $x2$), a simplified calculation of the harmonic mean can be calculated as:

$$HarmonicMean = \frac{2*x1*x2}{x1+x2}$$

As a harmonic mean penalizes extreme values, the F1-score uses a harmonic mean instead of averaging both values. In situations where precision is 1 and recall is 0, the harmonic mean is 0. If we were to use an average of both values, though, it would be 0.5, which is not suitable for a classifier model with high precision and low recall.

Mathematically, F1-score can be formulated as:

$$F1score = \frac{2 \cdot TruePositive}{2 \cdot TruePositive+FalsePositive+FalseNegative}$$

This can also be written as:

$$F1score = \frac{2 \cdot precision*recall}{precision+recall}$$

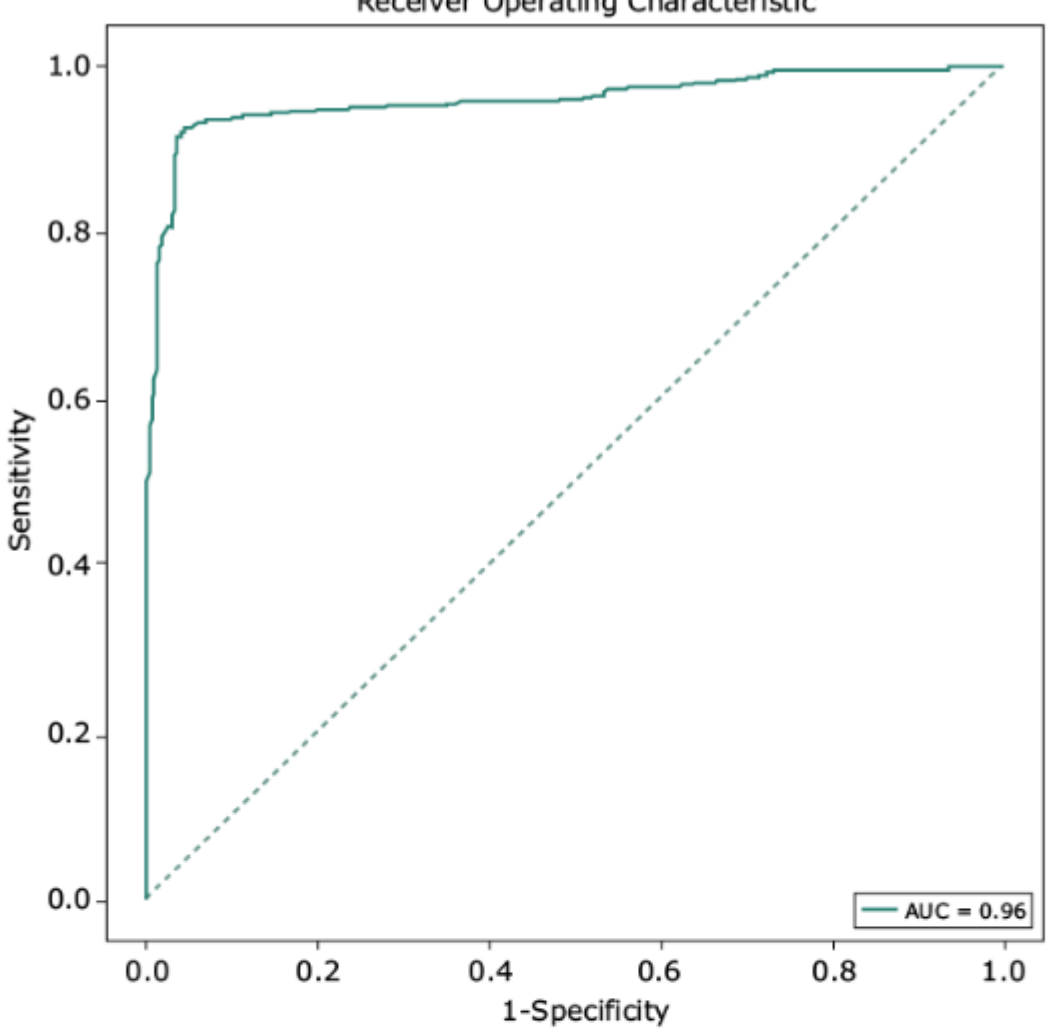


Figure 9.9. ROC curve

▶ Expand/collapse Extended Description

Area under the curve (AUC)

A receiver operating characteristic (ROC) curve can be used to evaluate the performance of a classifier. It is a plot where (1 – specificity) is plotted on the *x*-axis and sensitivity is plotted on the *y*-axis. ROC curve, as shown in figure 9.9, is used to understand how well the classifier can separate positive and negative examples and to identify the best threshold for separating them.

Area under the curve (AUC) or area under the receiver operating characteristic curve (AUROC) is the area under the ROC curve. AUC is commonly used to compare the classification performance of multiple classifiers. An AUC score of 1 is the ideal classifier.

Thus, data analytics has enabled the analysis of large amounts of healthcare data from thousands of patients and healthcare providers. Recent years have witnessed a change in treatment models, with more access to data. Like any other sector, the healthcare industry is undergoing a transformation. New applications have developed, which include:

- Operational improvements like increasing the efficiency of hospital staffs by predicting the inflow and outflow of patients
- Instant patient care, including real-time alerting of patients

- Predictive diagnostics, or using analytics to predict the likelihood of a patient developing certain medical conditions.

Data analytics and machine learning has been widely adopted in healthcare domain but it has the potential to do much more. The applications of data analytics and machine learning in healthcare have the potential to improve patient care, enhance clinical outcomes, and optimize health business management. Analytics and machine learning may perform regular activities such as patient record management, claims processing, and clinical recording. The primary goal is to eliminate human labor in administrative, regulatory, and management tasks that can be performed more efficiently by machines.

Data analytics and machine learning has a lot of potential in the healthcare field, but it also comes with some challenges, such as the unavailability good quality data, building user-friendly analytic systems for non-technical users like clinicians, and data security. There are also some ethical concerns, such as the privacy and security of patient data and its accountability. Even though there are some problems with ML in healthcare, the benefits far outweigh the problems.

Chapter 9 Review Exercises

Match the evaluation metric with its description.

a. The average of the difference between the ground truth and the predicted values
b. The average of the squared difference between the target value and the value predicted by the regression model
c. An indicator of the model's performance when all classes are equally important, measured as the proportion of correctly predicted events to all predicted events
d. The ratio of positive samples that were correctly classified to all samples that were classified as positive (either correctly or incorrectly); measures how well the model classifies a sample as positive

1. _____ Mean squared error (MSE)
2. _____ Mean absolute error (MAE)
3. _____ Accuracy
4. _____ Precision

Indicate whether each of the following is true or false.

5. Prescriptive analytics makes predictions about future outcomes using historical data
6. Data normalization is a step in data preprocessing.
7. Regression and classification are the two types of machine learning.
8. A data set is split into training and testing data to clean the data set.
9. Accuracy is a good test.
10. The AUROC score is not a good metric to compare the performance of classifiers.
11. Machine learning algorithms are usually trained on a data set called the training data set and tested on a data set called the transfer data set.
12. Logistic regression is a classification algorithm.
13. Precision, recall, and F1-score can be used as performance metrics for a regression problem.
14. Both supervised and unsupervised learning require at least one feature or attribute for training the algorithm.

Discussion

15. Describe four types of analytics, with one of their applications.

16. How can you handle missing data in a data set?
17. What is principal component analysis (PCA)? When do you use it?
18. Compare and contrast supervised and unsupervised learning.
19. Describe an appropriate performance metric for each of the following:
    a. Regression
    b. Classification
    c. Clustering

REFERENCES

Bas¸tanlar, Y., and M. Özuysal. 2014. Introduction to Machine Learning. Chapter 7 in ***miRNomics: MicroRNA Biology and Computational Analysis***, ed. M. Yousef and J. Allmer, 105–128. New York: Humana.

Bus¸oniu, L., R. Babuška, and B. de Schutter. 2010. Multi-Agent Reinforcement Learning: An Overview. Chapter 7 in ***Innovations in Multi-Agent Systems and Applications 1***, ed. D. Srinivasan and L. C. Jain, 183–221. New York: Springer.

Davenport, T. H. 2013. Analytics 3.0. ***Harvard Business Review*** 91(12):64–72.

Fang, H. 2015. Managing Data Lakes in Big Data Era: What's a Data Lake and Why Has It Became Popular in Data Management Ecosystem. In ***2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)***, 820–824. New York: IEEE.

García, S., J. Luengo, and F. Herrera. 2015. ***Data Preprocessing in Data Mining***. New York: Springer.

Goodfellow, I., Y. Bengio, and A. Courville. 2016. ***Deep Learning***. Cambrideg, MA: MIT Press.

Haneuse, S., D. Arterburn, and M. J. Daniels. 2021. Assessing missing data assumptions in EHR-based studies: A complex and underappreciated task. ***JAMA Network Open*** 4(2):e210184.

Hersh, W. R., A. M. Cohen, M. M. Nguyen, K. L. Bensching, and T. G. Deloughery. 2022. Clinical study applying machine learning to detect a rare disease: Results and lessons learned. ***JAMIA Open*** 5(2):ooac053.

Jiao, Y., and P. Du. 2016. Performance measures in evaluating machine learning based bioinformatics predictors for classifications. ***Quantitative Biology*** 4(4):320–330.

Norori, N., Q. Hu, F. M. Aellen, F. D. Faraci, and A. Tzovara. 2021. Addressing bias in big data and AI for health care: A call for open science. ***Patterns*** 2(10):100347.

Zheng, A., and A. Casari. 2018. ***Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists***. Sebastopol, CA: O'Reilly Media.

Chapter

10

Privacy of Health Information

By Ming Zhang, PhD, MS and Laurie Rinehart-Thompson, JD, RHIA, CHP, FAHIMA

Learning Objectives

- Interpret the HIPAA Privacy Rule and regulations
- Identify protected health information and entities covered by the HIPAA Privacy Rule
- Distinguish authorization requirements for uses and disclosures of protected health information
- Describe rights granted to individuals per the HIPAA Privacy Rule
- Analyze the scope of the Information Blocking Rule, including its eight exceptions
- Relate administrative requirements imposed by the HIPAA Privacy Rule
- Analyze breach notification requirements
- Address additional legal considerations relevant to HIPAA

KEY TERMS

Access
Accounting of disclosures
Administrative simplification
American Recovery and Reinvestment Act (ARRA)
Authorization
Business associate (BA)
Business associate agreement (BAA)
Civil monetary penalties (CMPs)
Covered entity
Deidentified health information
Designated record set

E-discovery
Electronic media
Electronic protected health information (ePHI)

Enforcement Rule
Health Information Technology for Economic and Clinical Health (HITECH) Act
Health Insurance Portability and Accountability Act of 1996 (HIPAA)
Health plan
Healthcare clearinghouse
Healthcare provider
Legal hold
Limited data set
Minimum necessary
Mitigation
Notice of Privacy Practices (NPP)
Office for Civil Rights
Operations
Payment
Pre-empt
Protected health information (PHI)
Resolution agreements
Transaction
Treatment
Workforce

Throughout history, healthcare consumers have been concerned about the privacy of their health information, which is essential for quality patient care, facilitating personhood, and preventing health data misuse and economic harm. In general, most patients in the United States understand that the communications between patients and their physicians are confidential. However, beyond this, most patients would probably agree they have little understanding about their rights to access their health information, who else has a legal right to the information, and how it may be used. State laws, which vary significantly, have provided some privacy protections, but nuances exist that challenge both patients and the healthcare professionals who provide treatment. Because of the need for federal privacy standards and greater efficiencies in the US healthcare system, industry leaders have worked on solutions to these challenges for decades.

In an effort to improve the efficiency and effectiveness of the nation's healthcare system, the **Health Insurance Portability and Accountability Act of 1996 (HIPAA)** was enacted by the US Congress and signed into law by President Bill Clinton (HIPAA 1996). HIPAA is a complex federal law that addresses the provision of continuity of health coverage, the control of fraud and abuse in healthcare, the reduction of healthcare costs, and the security and **administrative simplification** that includes (1) adopting standards for transaction and code sets used to exchange health data, (2) standard identifiers for health plans, providers, employees, and individuals used in standard transactions, and (3) security and privacy of Protected Health Information (PHI) (45 CFR 160, 162 164). For the purposes of this chapter, the focus is on the Administrative Simplification section of Title II, and even more specifically, the privacy and security of health information. HIPAA was updated in 2009 with the passage of American Recovery and Reinvestment Act (ARRA), which included the Health Information Technology for Economic and Clinical Health (HITECH) legislation, the issuance of the 2013 HIPAA

Omnibus Rule, and the guidance for patient access to their health information issued in 2016. All are addressed in this chapter.

HIPAA Privacy Rule

The **Office for Civil Rights** (OCR) of the U.S. Department of Health and Human Services (HHS) has been given responsibilities for the oversight and enforcement of the HIPAA regulations.

Since the Privacy Rule was originally issued, there have been many changes in the healthcare industry as well as technology so that the HIPAA Privacy Rule now includes regulations to include the HIPAA Breach Notification Rule; to modify the HIPAA Privacy Rule to include the provisions of section 105 of Title I of the Genetic Information Nondiscrimination Act (GINA) of 2008; and to try to increase the workability of the HIPAA regulations for covered entities (CEs) and business associates (BAs) (OCR 2013; Rockwern et al. 2021). Overall, the expanded Privacy Rule aims to strengthen the protection of health information from "unauthorized, deceptive, or harmful uses and align with the principles of medical ethics, respect individual rights, and support the culture of trust necessary to maintain and improve care delivery" (Rockwern et al. 2021).

Focus of HIPAA Privacy Regulations

The administrative simplification provisions within HIPAA focus on the adoption of standards for the electronic transmission of financial and administrative transactions, national identifiers, privacy, and security. With the expectation of reducing costs and administrative burdens, Congress also recognized that advances in electronic technology could erode the privacy of a person's health information. Taking this into consideration, Congress incorporated HIPAA legislation mandates for the adoption of federal privacy protections for individually identifiable health information. The HIPAA privacy regulations address the use and disclosure of **protected health information (PHI),** which is defined as "individually identifiable health information that is transmitted by electronic media, maintained in electronic media, or transmitted or maintained in any other form or medium" (45 CFR 160.103). Health plans, healthcare clearinghouses, and healthcare providers had to comply with the HIPAA Privacy Rule by April 14, 2003, with the exception of small health plans, which were given an additional year.

In the years that followed the enactment of HIPAA, some privacy advocates criticized the federal legislation as a "paper tiger" (Belfort 2010). Examples of criticisms were that privacy and security rules do not apply to many organizations that routinely handle large amounts of health information and the potential sanctions are not sufficiently severe except in rare cases of criminal conduct (Belfort 2010). Congress responded to these issues with the passage of the **Health Information Technology for Economic and Clinical Health (HITECH) Act**, which was enacted as part of the **American Recovery and Reinvestment Act (ARRA)** of 2009. The HITECH Act strengthened enforcement of the rules promulgated under HIPAA. This chapter covers federal legislation on health information privacy originally enacted under HIPAA and subsequently expanded and strengthened with the passage of the HITECH Act, with the issuance of the HIPAA Omnibus Rule in 2013, and the "Individuals' Right Under HIPAA to Access Their Health Information" (45 CFR 164.524).

Basics of HIPAA Privacy Regulations

Although the HIPAA privacy regulations are only one part of the original law, they are very complex, especially after the various expansions and clarifications. To fully understand them, it is important to understand the definitions and other foundations on which the regulations are built.

Covered Entities and Workforce

All of the various HIPAA privacy regulations apply to covered entities and all BAs of covered entities. Under HIPAA, a **healthcare provider** is any person or organization who furnishes, bills, or is paid for healthcare in the normal course of business (45 CFR 160.103 2013). A **covered entity** (CE) is a healthcare provider, health plan, or healthcare clearinghouse that transmits health information in electronic form in connection with healthcare transactions (OCR 2013). Healthcare providers include persons such as doctors, dentists, nurses, and therapists, as well as entities such as hospitals, clinics, urgent care centers, pharmacies, and institutions that provide and bill for services (OCR 2013). A **health plan** is an individual or group plan that provides or pays the cost of medical care (45 CFR 160.103 2013). Examples of health plans include group health plans, health insurance issuers, health maintenance organizations (HMOs), Medicare Part A or Part B, and the Medicaid program. A **healthcare clearinghouse** is defined by HIPAA as "a public or private entity, including a billing service, repricing company, community health management information system, or community health information system, and 'value-added' networks and switches, which do either of the following functions:

1. Processes or facilitates the processing of health information received from another entity in a nonstandard format or containing nonstandard data content into standard data elements or a standard transaction.
2. Receives a standard transaction from another entity and processes or facilitates the processing of health information into nonstandard format or nonstandard data content for the receiving entity" (OCR 2013).

When applying HIPAA, CEs must also consider members of their workforce, which may not be limited only to employees. The Privacy Rule defines **workforce** as "employees, volunteers, trainees, and other persons whose conduct, in the performance of work for a covered entity (CE) or BA, is under the direct control of such covered entity or business associate, whether or not they are paid by the covered entity or business associate" (OCR 2013). An example of nonpaid workforce would be a health information technology (HIT) or health information management (HIM) student who is completing a professional practice experience in the hospital setting.

Business Associates

The term "business associate" has a very specific meaning for CEs when complying with HIPAA. **Business associates (BAs)** essentially are all entities that are not members of the workforce but are providing a service or performing a task on behalf of a CE that involves the use or disclosure of PHI. The full definition can be found in Figure 10.1.

Examples of typical BAs of the health information department are an outsourced medical transcription company and coding or reimbursement consultants. The individuals performing these functions are not members of the facility's workforce, but they use or disclose health information to perform their respective duties on behalf of the CE. The newly expanded definition explicitly includes health information organizations that will be involved in exchanging information electronically. The new definition also carries the BA obligation out to BA subcontractors so that everyone or anyone handling PHI on behalf of a CE should consider themselves a BA.

Figure 10.1. Definition of HIPAA business associate

Business associate:

1. Except as provided in paragraph (4) of this definition, business associate means, with respect to a covered entity, a person who:
   i. On behalf of such covered entity or of an organized health care arrangement (as defined in this section) in which the covered entity participates, but other than in the capacity of a member of the workforce of such covered entity or arrangement, creates, receives, maintains, or transmits protected health information for a function or activity regulated by this subchapter, including claims processing or administration, data analysis, processing or administration, utilization review, quality assurance, patient safety activities listed at 42 CFR 3.20, billing, benefit management, practice management, and repricing; or
   ii. Provides, other than in the capacity of a member of the workforce of such covered entity, legal, actuarial, accounting, consulting, data aggregation (as defined in § 164.501 of this subchapter), management, administrative, accreditation, or financial services to or for such covered entity, or to or for an organized health care arrangement in which the covered entity participates, where the provision of the service involves the disclosure of protected health information from such covered entity or arrangement, or from another business associate of such covered entity or arrangement, to the person.
2. A covered entity may be a business associate of another covered entity.
3. *Business associate* includes:
   i. A Health Information Organization, E-prescribing Gateway, or other person that provides data transmission services with respect to protected health information to a covered entity and that requires access on a routine basis to such protected health information.

Source: 45 CFR 160.103

Once a person or organization is identified as a BA, the CE must initiate a **business associate agreement (BAA)** that meets all HIPAA requirements, including any recent updates or changes. CEs should enlist the help of legal counsel to ensure that the BAA clearly documents the requirements of the BA "to implement administrative, physical, and technical safeguards that reasonably and appropriately protect the confidentiality, integrity, and availability of the PHI that it creates, receives, maintains, or transmits on behalf of the covered entity" (45 CFR 164.308(b)(1) and 45 CFR 164.314). As amended by HITECH, "a contract between the covered entity and a business associate must establish the permitted and required uses and disclosures of protected health information by the business associate" and provide specific content requirements of the agreement (HHS 2013). The contract may not authorize the BA to use or further disclose the information in a manner that would violate the requirements of HIPAA and requires termination of the contract if the CE or BA is aware of noncompliant activities of the other (45 CFR 164.504). Other areas to be addressed in the agreement include the role of agents or subcontractors who are provided PHI; safeguards to protect the PHI; procedures for the reporting of security incidents of which it becomes aware; and authority of the CE

to terminate the contract if the BA violates a material term of the contract. With an acceptable BAA in place, CEs may lawfully disclose PHI to BAs to perform services for the CE.

Before HITECH came into force, BAs who failed to properly protect patient information were liable to the CEs via their BAA but did not face governmental penalties. The HITECH Act expands HIPAA data privacy and security requirements to include the BAs of CEs. Under the HITECH Act, a person or organization identified as a BA, even without a BAA, is now directly subject to the following HIPAA privacy requirements:

- Impermissible uses and disclosures
- Failure to provide breach notification to the CE
- Failure to provide access to a copy of electronic PHI either to the CE or individual
- Failure to disclose PHI to the Secretary as required to investigate or determine the BA's compliance
- Failure to provide an accounting of disclosures (OCR 2013)

BAs are subjected to these requirements to ensure there will be appropriate safeguards to protect health information. If an entity does not meet the definition of a CE or BA, it does not have to comply with the HIPAA Rules (OCR 2017).

Check Your Understanding 10.1

Answer the following questions.

1. This federal agency is charged with responsibility for the oversight and enforcement of the HIPAA privacy regulations.
   a. The Office of the National Coordinator for Health Information Technology
   b. The Centers for Medicare and Medicaid Services
   c. The HHS Office for Civil Rights
   d. The HHS Office of the Inspector General
2. Health plans, health providers and healthcare clearinghouses are classified as this under the HIPAA Privacy Rule.
   a. Covered entities
   b. Business associates
   c. Consultants
   d. Workforce
3. Employees, volunteers, trainees, and other persons performing functions on behalf of covered entities and business associates, whether paid or not, are considered to be:
   a. Consultants
   b. Workforce
   c. Immune from prosecution under HIPAA
   d. None of the above
4. This piece of legislation makes business associates subject to all of the HIPAA privacy requirements, including governmental sanctions and penalties.
   a. Health Insurance Portability and Accountability Act
   b. Affordable Care Act
   c. 2013 HIPAA Omnibus Rule
   d. Health Information Technology for Economic and Clinical Health Act

Match the term with the correct definition.

a. Business associate

b. Covered entity
c. Healthcare clearinghouse
d. Protected health information
e. Transaction

5. _____ Must be individually identifiable health information to meet this definition
6. _____ Entity that processes or facilitates the processing of health information received from another facility or receives a standard transaction from another entity for processing
7. _____ Healthcare provider that conducts electronic health information transactions, health plan, or healthcare clearinghouse
8. _____ Transmission of information between two parties for financial or administrative healthcare activities
9. _____ Providing a service or function on behalf of a covered entity that Involves the use of disclosure of protected health Information

---

Protected Health Information

To better understand the HIPAA Privacy Rule and the goal of ensuring individuals' protection of health information, it is important to examine the meaning of PHI because this fundamental term is used throughout much of the legislation. As mentioned earlier in the chapter, PHI is defined as "individually identifiable health information that is transmitted by electronic media, maintained in electronic media, or transmitted or maintained in any other form or medium" (45 CFR 160.103), which is usually collected by a healthcare professional to identify an individual and determine appropriate healthcare. PHI includes, but is not limited to, demographic data, medical history and conditions, and insurance information. A three-part test to determine whether or not information meets the definition of PHI is as follows:

1. The information must either identify the person or provide a reasonable basis to believe the person could be identified from the information given.
2. The information must relate to one's past, present, or future physical or mental health condition; the provision of healthcare; or payment for the provision of healthcare.
3. The information must be held or transmitted by a CE or its BA (Brodnik et al. 2023).

Furthermore, the Privacy Rule provides guidance on what is not PHI and specifically excludes personnel and educational records. Health information noted in education records is covered by the Family Educational Rights and Privacy Act (FERPA, 20 USC 1232g). Employment records held by a CE in its role as employer are exempt from the Privacy Rule (OCR 2013). As an example, employee physical examination reports contained within personnel files would not be subject to the Privacy Rule (Brodnik et al. 2023).

Another point to consider when interpreting the Privacy Rule is that the information must be created, transmitted or received by a CE (or BAs). Therefore, information that patients communicate to their doctors is covered, but information that patients communicate to parties who are not CEs or BAs is not protected under HIPAA. An example of this would be a patient disclosing health information to a family member or friends. In this instance, HIPAA does not prevent the family or friends from disclosing the health information to others (Sullivan and Hartsfield 2020).

Electronic Protected Health Information

When looking at the privacy and security of patient health information, **electronic protected health information (ePHI)** must be considered to ensure that ePHI is properly stored and used. The term ePHI refers to "individually identifiable health information that is transmitted by electronic media or maintained in electronic media" (45 CFR 160.103). In other words, ePHI is one kind of electronic data that a medical professional collects and stores to determine and provide proper patient care. The HIPAA Security Rule, discussed in detail in Chapter 11, is a key law written to protect ePHI and to provide guidance for how electronic health information can be accessed appropriately (Brodnik et al. 2023).

page 263

HIPAA sets forth clear definitions of **electronic media**, which is helpful for interpreting the intent of the legislation. Electronic media is defined as

1. Electronic storage material on which data are or may be recorded electronically, including, for example, devices in computers (hard drives) and any removable/transportable digital memory medium, such as magnetic tape or disk, optical disk, or digital memory card;
2. Transmission media used to exchange information already in electronic storage media. Transmission media include, for example, the internet, extranet or intranet, leased lines, dial-up lines, private networks, and the physical movement of removable/transportable electronic storage media. Certain transmissions, including of paper, via facsimile, and of voice, via telephone, are not considered to be transmissions via electronic media if the information being exchanged did not exist in electronic form immediately before the transmission. (OCR 2013)

These definitions, along with the Privacy Rule and Security Rule, reinforce that any PHI that is electronically stored, or electronically stored information that is then transmitted, must be protected from inappropriate access. Ensuring the security of ePHI and maintaining compliance with HIPAA are increasingly important because more healthcare providers opt to use online documentation systems and, as a result, must take cybersecurity into consideration.

Deidentification of Protected Health Information

The Privacy Rule does not restrict the use or disclosure of **deidentified health information**. In the Privacy Rule summary, HHS explains that

> deidentified health information neither identifies nor provides a reasonable basis to identify an individual. There are two ways to deidentify information; either: (1) a formal determination by a qualified statistician; or (2) the removal of specified identifiers of the individual and of the individual's relatives, household members, and employers is required, and is adequate only if the covered entity has no actual knowledge that the remaining information could be used to identify the individual. (45 CFR 164.514(a))

45 CFR 164.514(b) provides a distinct list of data elements that a CE can remove to ensure that a patient's information is deidentified. The 18 data elements are listed in table 10.1. Deidentified health information is useful in areas such as research, decision support, and education in health professions to support the secondary use of data for comparative effectiveness studies and other endeavors.

page 264

page 265

Table 10.1 . Safe harbor method: deidentification of PHI

To meet the intent of HIPAA requirements the following identifiers of the individual or of relatives, employers, or household members of the individual, must be removed:	
1	Names
2	All geographic subdivisions smaller than a State, including street address, city, county, precinct, zip code, and their equivalent geocodes, except for the initial three digits of a zip code if, according to the current publicly available data from the Bureau of the Census: (1) The geographic unit formed by combining all zip codes with the same three initial digits contains more than 20,000 people; and (2) The initial three digits of a zip code for all such geographic units containing 20,000 or fewer people is changed to 000
3	All elements of dates (except year) for dates directly related to an individual, including birth date, admission date, discharge date, date of death; and all ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements may be aggregated into a single category of age 90 or older
4	Telephone numbers
5	Fax numbers
6	Electronic mail addresses
7	Social security numbers
8	Medical record numbers
9	Health plan beneficiary numbers
10	Account numbers
11	Certificate/license numbers
12	Vehicle identifiers and serial numbers, including license plate numbers
13	Device identifiers and serial numbers
14	Web Universal Resource Locators (URLs)
15	Internet Protocol (IP) address numbers
16	Biometric identifiers, including finger and voice prints
17	Full face photographic images and any comparable images
18	Any other unique identifying number, characteristic, or code except a code used for Reidentification purposes

Source: 45 CFR 164.514(b) 2013

In November 2012, the HHS OCR released guidance to ensure the deidentification of PHI meets the intent of the HIPAA Privacy Rule. Figure 10.2 is a graphic depicting the two acceptable methods. The first method requires an expert to use statistical or scientific methods to ensure there is a very small probability of any identification of the individuals. The second method, the safe harbor method, requires that the 18 data elements listed in table 10.1 be removed. However, the second method also requires that there be no actual knowledge that the residual data can identify an individual (OCR 2022a). For example, if the data were deidentified by the second method, yet retained a rare diagnosis code for a very small population of people, it is possible the person might be able to be recognized, so the data would not be "deidentified" according to HIPAA.

Reidentification

A CE is allowed to have a policy and procedure in place for reidentifying information that had been previously deidentified, specifically assigning a unique code to the set of de-identified health information to permit re-identification by the CE (OCR 2022a). This code or other means of record identification can be used provided that the code is not derived from information about the individual and cannot be translated in some manner to identify the individual. Furthermore, the CE cannot use or disclose the code for any other purpose and must not divulge the mechanism for reidentification (45 CFR 164.514). Disclosure of the code or other means of de-identified information to be re-identified is considered a disclosure of PHI (OCR 2022a).

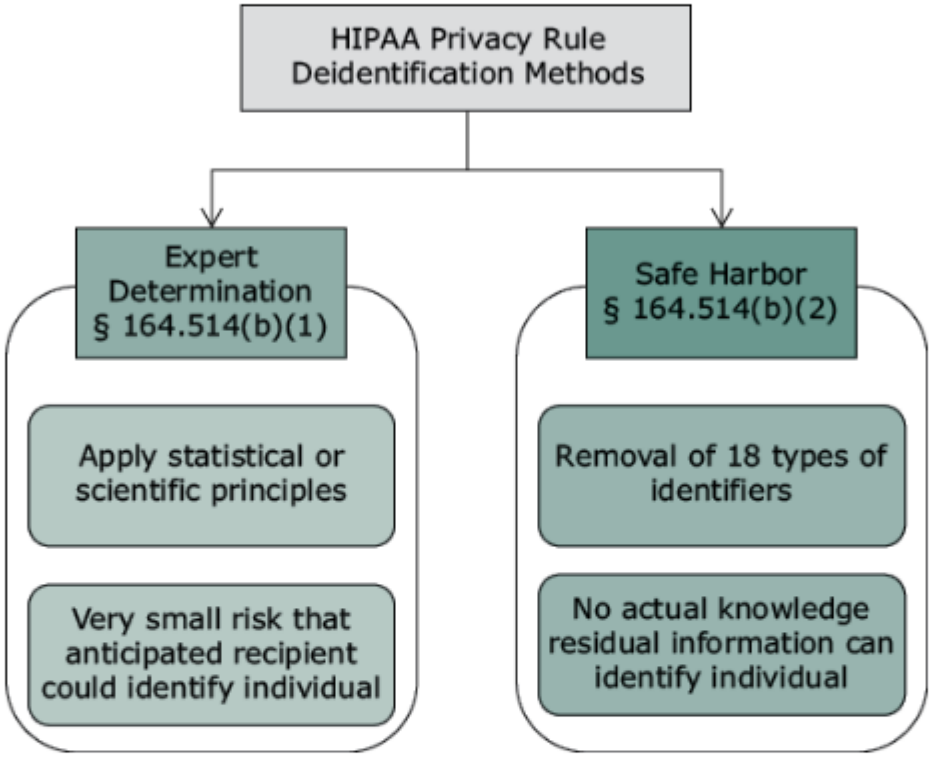


Figure 10.2. Two methods of HIPAA privacy rule deidentification

▶ Expand/collapse Extended Description

Source: OCR 2022

Table 10.2 . Healthcare transactions

1	Healthcare claims or equivalent encounter information
2	Healthcare payment and remittance advice
3	Coordination of benefits
4	Healthcare claim status
5	Enrollment and disenrollment in a health plan
6	Eligibility for a health plan
7	Health plan premium payments
8	Referral certification and authorization
9	First report of injury

10	Health claims attachments
11	Healthcare electronic funds transfers (EFT) and remittance advice
12	Other transactions the Secretary may prescribe by regulation

Source: AHIMA 2017

Healthcare Transactions

The HIPAA law amended by HITECH delineates the meaning of healthcare transactions. **Transaction** means "the transmission of information between two parties to carry out financial or administrative activities related to healthcare" (OCR 2013). This definition applies to transmissions of electronic health information. Table 10.2 provides a list of healthcare information transactions affected by HIPAA.

Minimum Necessary Requirement

To be in compliance with HIPAA, a CE must have policies and procedures in place that limit the protected information to be used or disclosed only to the amount necessary to achieve the purpose of the disclosure, the **minimum necessary**. This requires reasonable efforts to see, disclose, and request only the minimum amount of protected health information needed to accomplish the intended purpose of the use, disclosure, or request (45 CFR 164.502(b)).

Exceptions to the minimum necessary requirement are:

- Disclosures to or requests by a healthcare provider for treatment
- Uses or disclosures made to the individual
- Uses or disclosures made pursuant to an authorization
- Disclosures made to HHS for complaint investigation, compliance review or enforcement
- Uses or disclosures required by law
- Uses or disclosures required for compliance with other applicable requirements of HIPAA Transactions Rule or other HIPAA administrative simplification rules (45 CFR 164.502).

The minimum necessary information provided should be enough to address the intended purpose of the request but should also limit unnecessary or inappropriate information from being provided. The requirement was written with the minimum necessary words to allow it to be flexible to cover the many various circumstances associated with the request. The intent of updates to the HIPAA law was to provide more protection for the PHI while granting a greater level of control to the individual. This is in great contrast to the "any and all information" requests that were received in the past.

Uses and Disclosures of PHI

It is important to distinguish between the concepts of use and disclosure as emphasized in the HIPAA Privacy Rule:

- "Use means, with respect to individually identifiable health information, the sharing, employment, application, utilization, examination, or analysis of such information within an entity that maintains such information" (45 CFR 160.103).
- "Disclosure means the release, transfer, provision of access to, or divulging in any other manner of information outside the entity holding the information" (OCR 2013).

Authorization

An **authorization** is a written permission for a specific use or disclosure. As a general rule, authorization by the patient or the patient's designated representative is required for the use or disclosure of PHI unless it meets an exception where authorization is not required (Brodnik et al. 2023). This section examines the elements required for a valid authorization as well as uses and disclosures that are permitted without an authorization.

Elements of a Valid Authorization

A valid authorization must contain the following elements:

1. A description of the PHI to be used or disclosed;
2. The person authorized to make the use or disclosure;
3. The person to whom the covered entity may make the requested use or disclosure;
4. Expiration date or event; and
5. In some cases, the purpose for the use or disclosure of the information (HHS 2016).
6. Signature and date by the patient or other legally authorized individual

Other core elements include a statement of right to revoke; a statement regarding the possibility of redisclosure; a statement specifying that treatment, payment, enrollment, and eligibility for benefits cannot be denied when the individual declines to sign an authorization; and a copy of the signed authorization is provided to the individual when authorization for the use and disclosure is sought by the covered entity. One purpose of the authorization is to ensure that the CE and the individual fully understand the confidential relationship boundaries and what information is to be disclosed.

Authorization Not Required

There are numerous situations where authorization is not required prior to a use or disclosure. These exceptions are noted as follows.

To the Individual

When PHI is disclosed to the individual who is the subject of the PHI, an authorization signed by the individual or his legal representative (if applicable) is not required.

Treatment, Payment and Operations (TPO)

Treatment is the "provision, coordination, or management of health care and related services by one or more health care providers, including coordination or management of health care by a health care provider with a third party; consultation between health care provider relating to a patient or the referral of a patient for health care from one health care provider to another" (AHIMA 2017). An example of treatment would be the freedom of healthcare providers from two separate organizations involved in an individual's care to discuss that individual's diagnosis, treatment, and prognosis.

Payment consists of "activities undertaken by a health plan to obtain premiums or to determine or fulfill its responsibility for coverage and provision of benefits under the health plan; or a health care provider or health plan to obtain or provide reimbursement for the provision of health care" (AHIMA 2017). Billing, claims management and collection, utilization review, and reviewing healthcare services for medical necessity, coverage, or to justify charges are examples of payment activities.

Healthcare **operations** are "certain administrative, financial, legal, and quality improvement activities of a covered entity that are necessary to run its business and to support the core functions of treatment and payment" (45 CFR 164.506). Examples of operations include quality assessment and improvement activities; competency assurance activities; conducting or arranging for medical reviews, audits, or legal services; specified

insurance functions such as underwriting, enrollment, and premium rating; business planning, development, management, and administration; and business management and general administrative activities of the entity (Sullivan and Hartsfield 2020).

page 269

Opportunity to Agree or Object

There must be a balance between the patients' right to control the use and disclosure of their health information and the need for it to still be available when and where necessary. There are two situations where patients must verbally indicate whether or not their information may be used or disclosed (i.e., they must agree or object), but a written authorization is not required if they agree. The first situation is inclusion in the facility directory, from which individuals who ask about the patient by name may be given limited information, and clergy may be given names of patients who share a religious affiliation with the clergy member. The second situation is disclosure of relatives or friends directly involved in the patient's care or payment, if communication with the patient is not possible. The CE must use its best judgment in these communications, especially when agreement cannot yet be obtained (i.e., emergency situations).

Incidental Use and Disclosure

There is a potential risk for an individual's health information to be disclosed incidentally due to established practices in healthcare communications and various environments. An example is in the emergency department where a patient or family member might overhear a discussion between providers regarding the patient in the next treatment room. A use or disclosure of this nature is permitted as long as the CE has adopted reasonable safeguards as required by the HIPAA Privacy Rule. Reasonable safeguards include speaking quietly when discussing a patient's condition with family members in a public area such as a waiting room and avoiding using patient names in public places. In addition, proactive measures such as training and posting signs to remind employees to protect patient confidentiality are also reasonable safeguards (OCR 2013).

Public Interest and Benefit Activities

The Privacy Rule permits the use and disclosure of patient information as it relates to national priority issues. Twelve public interest purposes are identified in the Privacy Law and must be reported when required by law, for public health activities, serious threats to health or safety, and research. All twelve public interest purposes contain special conditions that are outlined in 42 CFR 164.512. The public interest purposes are:

- Required by law: statute, regulation, court orders
- Public health activities: for preventing or controlling disease, subject to FDA regulation, cases of communicable disease, and work-related illness or injury and surveillance
- Victims of abuse, neglect, or domestic violence when necessary
- Health oversight activities: for legally authorized activities including audits and investigations related to the healthcare system and governmental benefit programs

page 270

- Judicial and administrative proceedings: with appropriate notice to the individual or protective order provided
- Law enforcement purposes: with conditions that must be met for the disclosure
- Decedents: for identifying or determining the cause of death
- Cadaveric organ, eye, or tissue donation: to facilitate donation and transplantation
- Research: when privacy rule definition is met and in specific instances
- Serious threat to health or safety: to prevent or lessen a serious imminent threat to person or public

- Essential government functions: those authorized by law for proper execution of a military mission or conducting intelligence and national security activities
- Worker's compensation: as allowed by applicable laws (42 CFR 164.512)

Limited Data Set

A **limited data set** is "PHI that excludes direct identifiers of the individual and the individual's relatives, employers, or household members but still does not deidentify the information" (AHIMA 2017). Limited data sets may be used for research, public health, or healthcare operations, provided that the recipient of the data set enters into a data use agreement that requires the recipient to safeguard the information.

Individual Rights Regarding Health Information

The HIPAA Privacy Rule gives individuals many rights with respect to their health information. They are outlined in this section.

Access to PHI

45 CFR 164.524 addresses an individual's right of **access** to "inspect and obtain a copy of protected health information about the individual in a designated record set, for as long as the protected health information is maintained in the designated record set." The **designated record set** contains the "medical records and billing records about the individual and the enrollment, payment, claims adjudication, and case or medical management record systems and used for the CE to make decisions about the individual" (45 CFR 164.501). Exceptions to the right to access include psychotherapy notes; information compiled in reasonable anticipation of a civil, criminal, or administrative action or proceeding; PHI subject to the Clinical Laboratory Improvements Amendments of 1988, 42 USC § 263a; and PHI that is exempt from the Clinical Laboratory Improvements Amendments of 1988, pursuant to 42 CFR 493.3(a)(2) (45 CFR 164.524).

When a request is received for access to PHI, the CE must respond in a timely manner for both paper and electronic records. Individuals have the right to specify the form and format for the CE to provide access to their PHI, if it can be readily producible as requested or a copy can be provided in another form and format agreed on by both parties. If a paper copy is requested, the information should be provided in paper form whether the original is in paper or electronic form. When the request is for an electronic copy of a paper record, the CE is required to provide an electronic copy if it can be readily producible in that format, such as by scanning the paper record into an electronic format. If an electronic copy of electronically stored information is requested, the CE is required to provide access if this can be accomplished in a reasonable manner and in an agreed-on alternative electronic format (CFR 164.524). When patients seek to obtain a copy of their own records, the HIPAA Privacy Rule limits fees to those that are reasonable and cost-based (Brodnik et al. 2023).

Denial of Access

Access to PHI may be denied with the right, in some cases, for the denial to be reviewed by a licensed professional who was not a part of the original determination of the request. Certain limited explicit situations are considered to be unreviewable, which means that the denial remains with no further review of the request.

- For psychotherapy notes or information compiled for potential legal proceedings.

- By an inmate for a copy of PHI where the CE is a correctional institution, or provider acting at direction of the institution, and the copy would jeopardize the health, safety, security, custody, or rehabilitation of the inmate or other inmates, or safety of correctional officers, employees, or others at the institution or responsible for transporting the inmate. The inmate retains the right to inspect the PHI.
- That are in a designated record set part of a research study still in progress that includes treatment (for example, clinical trial) provided the consent to participate in the study required the individual to agree to a temporary suspension of access. The right of access is reinstated when the study is completed.
- For PHI in records protected by the federal Privacy Act such as those under the control of a federal agency, if the denial is consistent with the Privacy Act requirements.
- Where the requested information was obtained by someone other than a healthcare provider such as a family member of the individual who was assured confidentiality, and the source of the information would reasonably be revealed if access granted (OCR 2013).

Right to Authorize Access to PHI

The person who consented to the treatment, typically an adult patient who is mentally and physically competent, can also authorize access to the PHI. Other parties who can authorize access include

- A parent, guardian, or custodian of a minor patient under 18 years of age.
- A parent, guardian, or custodian of an incompetent patient.
- Legal healthcare representative.
- Power of attorney for healthcare.
- Personal representative or executor or administrator of a deceased patient's estate.

The other parties mentioned here are considered to be the personal representative for the patient and are authorized to make healthcare decisions.

Amendment Requests

Individuals may request a CE to amend PHI or a record in a DRS (45 CFR 164.526). The CE may deny the request if it determines that the PHI or the record

- Was not created by the CE
- Is not part of the DRS
- Is not available for inspection
- Is accurate and complete

A CE may require that an **amendment request** be presented in writing, with the reason, if the individual knows this requirement ahead of time. Individuals are usually notified via the CE's Notice of Privacy Practices (NPP).

The CE must respond within 60 days by allowing or denying the amendment in writing. Further, it may extend its response by 30 days as long as it explains in writing why there will be a delay and a date to expect a response.

For granted requests, the CE must:

1. Identify the affected records in the DRS append them by linking them to the location of the amendment.

2. Tell the individual the amendment was accepted, giving him or her the opportunity to identify who the amendment is to be shared with. The individual's agreement to share must also be obtained, and sharing must be completed within a reasonable amount of time based on reasonable efforts to those who previously received the information.

Amendment denials must be provided within 60 days and written in plain language. They must contain:

- The basis for the denial
- The individual's right to submit a written statement disagreeing with the denial
- The process by which the individual can submit his or her disagreement
- A statement explaining how, if the individual does not submit a disagreement, he or she may request that both the original amendment request and the CE's denial accompany any future disclosures of the PHI in question
- A description of how the individual may complain to the CE, providing contact information (45 CFR 164.526(d))

The CE can prepare a written rebuttal to an individual's disagreement with a denial. A copy of the rebuttal must be given to the individual. The following must be appended or linked to the subject PHI or record: requests for amendments, denials, the individual's disagreement (if one exists), and the CE's rebuttal, if one was created. In future disclosures of the information, the material or a summary must accompany them. If a request for amendment was denied and no disagreement was submitted by the individual, the request for amendment and denial must only accompany future disclosures upon individual request (Brodnik et al. 2023).

Accounting of Disclosures

Individuals have a right to an **accounting of disclosures**, so they know to whom their PHI has been disclosed. Not all disclosures must occur in an accounting. Exceptions are:

1. For treatment, payment, or healthcare operations unless disclosure occurs through an EHR
2. PHI given to individual
3. Incidental to an otherwise permitted or required use or disclosure
4. Pursuant to an authorization
5. For use in a facility directory, to notify others involved in the individual's care, or for other notification purposes
6. To meet national security or intelligence purposes
7. To correctional institutions or law enforcement officials
8. Disclosure made as part of a limited data set
9. Disclosure that occurred before the HIPAA privacy compliance date (45 CFR 164.528(a)(1))

The accounting report must include the date of disclosure, name and address of the recipient of the information, and a statement of the purpose of the disclosure (or a copy of a written request for disclosure) (45 CFR 164.528(b)).

The accounting must include non-exempted disclosures made during the previous three years. The time limit for responding to a request for accounting is 60 days. It may be extended once by 30 days. The first annual accounting request by an individual is free,

and subsequent request may result in a reasonable, cost-based fee. Documentation regarding accounting of disclosures must be kept for six years (45 CFR 164.528(d)).

Confidential Communications Request

Healthcare providers and health plans are obligated to give individuals the right to request **confidential communications**. This means that it can be requested that communications of PHI be routed to an alternative location or by an alternative method (45 CFR 164.522(b)). Reasonable requests to healthcare providers do not require a reason. Health plans must honor reasonable requests if disclosure could present a safety risk. Requests may be refused if information is not provided regarding how payment will be handled or an alternative address or method of contact is not provided.

Individuals may request confidential communications to preclude family members from seeing information such as billing statements from a psychiatrist who is providing counseling for domestic violence. In this case, the individual may request that the bills be sent to a work address instead of the individual's home, where the abuser also lives (Brodnik et al. 2023).

Restriction Requests

An individual may **request restrictions** by asking a CE to restrict the uses and disclosures of PHI for carrying out TPO (45 CFR 164.522(a)). If a CE agrees to a restriction, it must abide by it. Either the CE or individual may terminate a restriction, although the CE must inform the individual of the termination.Termination only applies to PHI created or received after the individual has been informed (45 CFR 164.522(a)).

Generally the granting of restriction requests is optional. Unless otherwise required by law, however, one situation where a restriction request must be honored is if the disclosure would be made to a health plan for payment or operations (and not for treatment) and the PHI is relevant to an item or service paid for in full by a person or entity other than the health plan (Brodnik et al. 2023).

File Complaints

Individuals must have a way to complain about a CE's alleged violations of the Privacy Rule. The CE's NPP must contain contact information for the organization itself as well as how to submit complaints to the OCR at the US Department of Health and Human Services. All complaints and dispositions must be documented by the CE (Brodnik et al. 2023).

AHIMA is committed to the advocacy and protection of consumers with regard to their health information. The ***AHIMA Consumer Health Information Bill of Rights*** provides a model for protecting health information principles, as shown in Figure 10.3, and aligns with the rights provided by the HIPAA Privacy Rule.

Figure 10.3. AHIMA Consumer Health Information Bill of Rights

An individual has the right to:

1. Look at their health information and/or get a paper or electronic copy of it
2. Accurate and complete health information
3. Ask for changes to their health information
4. Know how their health information is used or shared and who receives it
5. Ask for limitations on the use and release of their health information
6. Expect their health information to be private and secure
7. Be informed about privacy and security breaches of their health information

8. File a complaint or report a violation regarding their health information

Source: AHIMA (2015)

Check Your Understanding 10.2

Answer the following questions.

1. Protected health information, subject to HIPAA protection, is defined by the following criteria except:
   a. It is identifiable health information regardless of who holds or uses it
   b. Identifies the person or provides a reasonable basis to believe the person could be identified from the information given
   c. Relates to one's past, present, or future physical or mental health condition, the provision of care or payment for the provision of healthcare
   d. The information must be held or transmitted by a covered entity or its business associate
2. The HIPAA methods titled Expert Determination and Safe Harbor are ways in which the following can be achieved legally.
   a. Data analysis
   b. Reidentification
   c. Deidentification
   d. Public health reporting
3. Protected health information which is maintained in a designated record set can be accessed by the patient or other authorized party upon request. Covered entities and business associates must respond to requests within the following time frame after receipt of the request.
   a. In a timely manner
   b. 60 days
   c. 30 days
   d. 15 days
4. If a patient has health insurance, but pays in full for a healthcare service themselves and asks that the information be kept private the covered entity must
   a. Release the information to the health insurance
   b. Get special patient consent to release the information
   c. Comply with the patient's request and keep the information private
   d. Request permission from HHS to release the information
5. There are instances under the HIPAA Privacy Rule where a person can informally agree or object to the release of their protected health information. All of the following are examples except:
   a. Inclusion of the patient name in a facility directory
   b. When releasing information for public health purposes
   c. When family and friends are with the patient being treated
   d. When releasing information for disaster relief

Information Blocking and the Cures Act

HIPAA has substantially advanced the ability of individuals to access their health information through its individual right of access (as discussed in the Individual Rights

section in this chapter) and, more recently, the Right of Access Initiative that has been aggressively enforced (OCR 2022b). However, the ability of patients to access their health information has been prioritized and dramatically expanded further via passage of the **21st Century Cures Act (Cures Act)**, which was signed into law on December 13, 2016. The Act's overarching purpose is to accelerate innovations in healthcare (FDA 2020) and this includes a simpler exchange of patient information among patients, providers and payers, giving patients greater access to their electronic health information and allowing them to access it through apps or devices that they choose.

Practically, the Cures Act has sought to normalize the sharing of patient information as a standard element of – not an exception to -- patient care (ONC 2022). The Cures Act specifically addresses this through the **Information Blocking Rule.** Designed to eliminate barriers that impede sharing and exchange of patient information, the Information Blocking Rule initially took effect on April 5, 2021 (Tripathi and Posnack 2021) and became fully operational on October 6, 2022, a date informally referred to as "data liberation day" because of the expansive scope of digital records that patients can now access. **Information blocking** is prohibiting conduct that "is likely to interfere with access, exchange, or use of **electronic health information (EHI)**." (45 CFR 171.103). EHI is electronic protected health information (ePHI) that would be included in a designated record set (DRS), regardless of whether the group of records is used or maintained by or for a CE or BA. Information that an individual has a legal right to per HIPAA is subject to information blocking. The EHI definition excludes psychotherapy notes and information compiled in reasonable anticipation of or for use in a civil, criminal, or administrative action or proceeding (45 CFR 171.102). It also does not include information covered by the Federal Educational Rights and Privacy Act (FERPA), employment records, records of an individual deceased greater than fifty years, or de-identified health information (ONC 2022).

The Office of the National Coordinator (ONC) provides further guidance for determining whether information is EHI. This is depicted in Figure 10.4, which is adapted from ONC materials. Figure 10.5 further describes the scope of EHI.

Information blocking prohibitions apply to **actors,** which include health care providers, health information technology developers of certified health IT, health information exchanges, and health information networks (45 CFR 171.102). In order not to interfere with access, exchange or use of EHI, actors must make the EHI available to patients or their designees without unnecessary or intentional delay. There are two separate standards for determining culpability, depending on who the actor is. The standard applied to health information technology developers, health information exchanges or health information networks is whether they knew or should have known that their actions were likely to interfere with access, exchange or use. The standard applied to a health care provider is whether the provider knew that their actions were likely to interfere with access, exchange or use of EHI (45 CFR 171.103).

Figure 10.4. Determination of EHI Status

Step 1	Is it individually identifiable health information that is maintained in electronic media or transmitted by electronic media? If yes, proceed to step 2.
Step 2	Would it be included in one of the following? Medical records and billing records of a provider about individuals Enrollment, payment, claims adjudication, and case or medical management record systems maintained by or for a health plan Records used in whole or in part, to make decisions about individuals If yes, proceed to step 3.
Step 3	Is it excluded from the EHI definition? If no, then proceed to step 4.
Step 4	It is EHI.

Source: ONC 2022

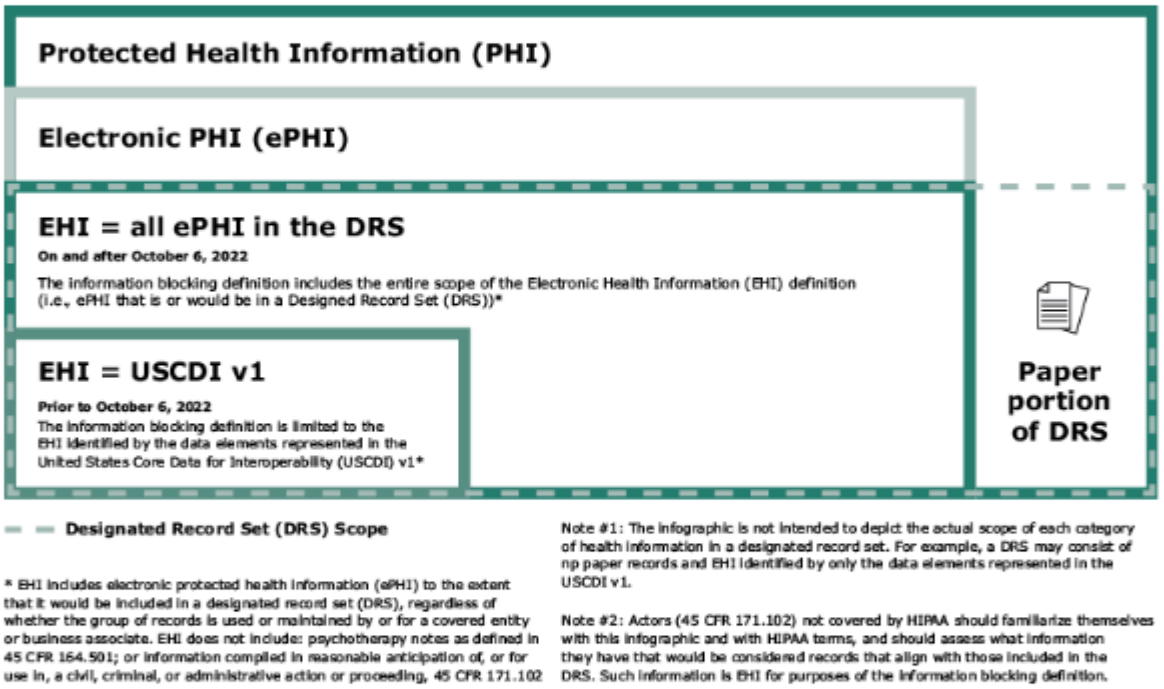

Protected Health Information (PHI)
Electronic PHI (ePHI)
EHI = all ePHI in the DRS
On and after October 6, 2022
The information blocking definition includes the entire scope of the Electronic Health Information (EHI) definition (i.e., ePHI that is or would be in a Designed Record Set (DRS))*
EHI = USCDI v1
Prior to October 6, 2022
The information blocking definition is limited to the EHI identified by the data elements represented in the United States Core Data for Interoperability (USCDI) v1*
Paper portion of DRS
Designated Record Set (DRS) Scope
* EHI includes electronic protected health information (ePHI) to the extent that it would be included in a designated record set (DRS), regardless of whether the group of records is used or maintained by or for a covered entity or business associate. EHI does not include: psychotherapy notes as defined in 45 CFR 164.501; or information compiled in reasonable anticipation of, or for use in, a civil, criminal, or administrative action or proceeding, 45 CFR 171.102
Note #1: The infographic is not intended to depict the actual scope of each category of health information in a designated record set. For example, a DRS may consist of np paper records and EHI identified by only the data elements represented in the USCDI v1.
Note #2: Actors (45 CFR 171.102) not covered by HIPAA should familiarize themselves with this infographic and with HIPAA terms, and should assess what information they have that would be considered records that align with those included in the DRS. Such information is EHI for purposes of the information blocking definition.

Figure 10.5. Scope of EHI

▶ Expand/collapse Extended Description

Source: OCR 2022

Not every action that appears to be information blocking is a violation. There are eight identified situations where it may be reasonable and necessary not to share or exchange information. These situations are exceptions to information blocking if certain conditions are met (ONC 2022) and they are divided into two categories: 1) exceptions that involve ***not*** fulfilling requests to access, exchange, or use EHI (five exceptions), and 2) exceptions that involve ***procedures*** for fulfilling requests to access, exchange, or use EHI (three exceptions). If information is withheld or delayed, but the action does not meet one of the eight exceptions, this does not automatically mean that information blocking has taken place. These practices will be evaluated individually (ONC 2022). The exceptions

are displayed in Figure 10.6. Figure 10.7, which is adapted from ONC materials, outlines the conditions for each exception to be recognized.

To facilitate efficient and secure data exchange, the Cures Act requires health IT developers to publish **APIs (application programming interfaces)** that allow different software applications to easily interact with one another and enable patients to access their EHI via the web and various applications (Glaser and Gardner 2022). The Cures Act specifies the use of **Fast Healthcare Interoperability Resource** (FHIR) to prescribe data formats. FHIR is strictly controlled by Health Level Seven International (HL7), an organization with significant corporate and organizational members, including health technology vendors, professional organizations, government agencies, and large health systems (Glaser and Gardner 2022).

Much of the attention on the Information Blocking Rule has centered around healthcare organizations' patient portals, although neither the Cures Act nor the Information Blocking Rule actually mentions patient portals (Dworkowitz 2022). Patient portals are where providers can change their practices the most significantly to foster patient access, and where patients have seen the greatest changes. For example, patient portals that once offered rather limited information and reports now include a much greater variety including operative reports, lab results, and physician progress notes that more completely document a patient's story (Dworkowitz 2022). With the ONC's interpretation of the rule to mean that patients should be able to obtain their information as soon as possible, the patient portal has become an obvious tool for ensuring that this happened (Dworkowitz 2022).

page 279

Figure 10.6. Exceptions to the information blocking provision

▶ Expand/collapse Extended Description

Source: ONC 2022

page 280

Figure 10.7. Conditions for Information Blocking Exceptions

Exception	Description	Conditions That Must Be Met
Exceptions that involve not fulfilling requests to access, exchange, or use EHI:		
Preventing Harm (1)	It will not be information blocking for an actor to engage in practices that are reasonable and necessary to prevent harm to a patient or another person, provided certain conditions are met.	*All* of the following bulleted points must be met. The actor must hold a reasonable belief that the practice will substantially reduce a risk of harm; • The actor's practice must be no broader than necessary; • The actor's practice must satisfy at least one condition from each of the following categories: type of risk, type of harm, and implementation basis; and • The practice must satisfy the condition concerning a patient right to request review of an individualized determination of risk of harm.

Privacy (2)	It will not be information blocking if an actor does not fulfill a request to access, exchange, or use EHI in order to protect an individual's privacy, provided certain conditions are met.	*At least one* of the numbered sub-exceptions must be met. 1. Precondition not satisfied: If an actor is required by a state or federal law to satisfy a precondition (such as a patient consent or authorization) prior to providing access, exchange, or use of EHI, the actor may choose not to provide access, exchange, or use of such EHI if the precondition has not been satisfied under certain circumstances. 2. Health IT developer of certified health IT not covered by HIPAA: If an actor is a health IT developer of certified health IT that is not required to comply with the HIPAA Privacy Rule, the actor may choose to interfere with the access, exchange, or use of EHI for a privacy-protective purpose if certain conditions are met. 3. Denial of an individual's request for their EHI consistent with 45 CFR 164.524(a) (1) and (2): An actor that is a covered entity or business associate may deny an individual's request for access to his or her EHI in the circumstances provided under 45 CFR 164.524(a)(1) and (2) of the HIPAA Privacy Rule. 4. Respecting an individual's request not to share information: An actor may choose not to provide access, exchange, or use of an individual's EHI if doing so fulfills the wishes of the individual, provided certain conditions are met.

Security (3)	It will not be information blocking for an actor to interfere with the access, exchange, or use of EHI in order to protect the security of EHI, provided certain conditions are met.	*All of the numbered* sub-exceptions and the *bulleted point* must be met. • The practice must be: 1. Directly related to safeguarding the confidentiality, integrity, and availability of EHI; 2. Tailored to specific security risks; and 3. Implemented in a consistent and non-discriminatory manner. • The practice must either implement a qualifying organizational security policy or implement a qualifying security determination.
Infeasibility (4)	It will not be information blocking if an actor does not fulfill a request to access, exchange, or use EHI due to the infeasibility of the request, provided certain conditions are met.	The practice must meet *one of the following conditions*: • Uncontrollable events: The actor cannot fulfill the request for access, exchange, or use of electronic health information due to a natural or human-made disaster, public health emergency, public safety incident, war, terrorist attack, civil insurrection, strike or other labor unrest, telecommunication or internet service interruption, or act of military, civil or regulatory authority. • Segmentation: The actor cannot fulfill the request for access, exchange, or use of EHI because the actor cannot unambiguously segment the requested EHI. • Infeasibility under the circumstances: The actor demonstrates through a contemporaneous written record or other documentation its consistent and non-discriminatory consideration of certain factors that led to its determination that complying with the request would be infeasible under the circumstances. • Additionally, the actor must provide a written response to the requestor within 10 business days of receipt of the request with the reason(s) why the request is infeasible.

Health IT Performance (5)	It will not be information blocking for an actor to take reasonable and necessary measures to make health IT temporarily unavailable or to degrade the health IT's performance for the benefit of the overall performance of the health IT, provided certain conditions are met.	*All of the numbered* sub-exceptions and the *bulleted* point must be met. The practice must: 1. Be implemented for a period of time no longer than necessary to achieve the maintenance or improvements for which the health IT was made unavailable or the health IT's performance degraded; 2. Be implemented in a consistent and non-discriminatory manner; and 3. Meet certain requirements if the unavailability or degradation is initiated by a health IT developer of certified health IT, HIE, or HIN. An actor may take action against a third-party app that is negatively impacting the health IT's performance, provided that the practice is: 1. For a period of time no longer than necessary to resolve any negative impacts; 2. Implemented in a consistent and non-discriminatory manner; and 3. Consistent with existing service level agreements, where applicable. • If the unavailability is in response to a risk of harm or security risk, the actor must only comply with the Preventing Harm or Security Exception, as applicable.
Exceptions that involve procedures for fulfilling requests to access, exchange, or use EHI:		

Content and Manner (6)	It will not be information blocking for an actor to limit the content of its response to a request to access, exchange, or use EHI or the manner in which it fulfills a request to access, exchange, or use EHI, provided certain conditions are met.	The points for each condition must be met. Content Condition: Establishes the content an actor must provide in response to a request to access, exchange, or use EHI in order to satisfy the exception. 1. Up to 24 months after the publication date of the Cures Act final rule, an actor must respond to a request to access, exchange, or use EHI with, at a minimum, the EHI identified by the data elements represented in the United States Core Data for Interoperability (USCDI) standard. 2. On and after 24 months after the publication date of the Cures Act final rule, an actor must respond to a request to access, exchange, or use EHI with EHI as defined in § 171.102. Manner Condition: Establishes the manner in which an actor must fulfill a request to access, exchange, or use EHI in order to satisfy this exception. • An actor may need to fulfill a request in an alternative manner when the actor is: ◦ Technically unable to fulfill the request in any manner requested; or ◦ Cannot reach agreeable terms with the requestor to fulfill the request. • If an actor fulfills a request in an alternative manner, such fulfillment must comply with the order of priority described in the manner condition and must satisfy the Fees Exception and Licensing Exception, as applicable

Fees (7)	It will not be information blocking for an actor to charge fees, including fees that result in a reasonable profit margin, for accessing, exchanging, or using EHI, provided certain conditions are met.	The points for each condition must be met. The practice must: Meet the basis for fees condition. • For instance, the fees an actor charges must: ◦ Be based on objective and verifiable criteria that are uniformly applied for all similarly situated classes of persons or entities and requests. ◦ Be reasonably related to the actor's costs of providing the type of access, exchange, or use of EHI. Not be based on whether the requestor or other person is a competitor, potential competitor, or will be using the EHI in a way that facilitates competition with the actor. • Not be specifically excluded. • For instance, the exception does not apply to: ◦ A fee based in any part on the electronic access by an individual, their personal representative, or another person or entity designated by the individual to access the individual's EHI. ◦ A fee to perform an export of electronic health information via the capability of health IT certified to § 170.315(b)(10). • Comply with Conditions of Certification in § 170.402(a)(4) (Assurances – certification to "EHI Export" criterion) or § 170.404 (API).

Licensing (8)	It will not be information blocking for an actor to license interoperability elements for EHI to be accessed, exchanged, or used, provided certain conditions are met.	All conditions must be met.
		The practice must meet:
		• The negotiating a license conditions: An actor must begin license negotiations with the requestor within 10 business days from receipt of the request and negotiate a license within 30 business days from receipt of the request.
		• The licensing conditions:
		  ◦ Scope of rights
		  ◦ Reasonable royalty
		  ◦ Non-discriminatory terms
		  ◦ Collateral terms
		  ◦ Non-disclosure agreement
		• Additional conditions relating to the provision of interoperability elements.

Source: ONC 2022.

Claims of information blocking can be reviewed by the ONC per its health IT certification program. The Department of Health and Human Services' Office of the Inspector General has the authority to investigate information blocking claims against all actors who are covered by the Information Blocking Rule (ONC 2022).

Check Your Understanding 10.3

1. Which of the following is an actor per the Information Blocking Rule?
   a. MedPlus, a health insurance company
   b. LifePlus, a life insurance company
   c. CliniMatters, an IT developer of certified EHRs
   d. CodeStat, a business associate of a healthcare system
2. A claim is filed against Community Health System, alleging that it violated the Information Blocking Rule by interfering with a patient's access to their EHI. In reviewing this case, the standard applied to Community Health System is:
   a. Whether it knew its actions were likely to interfere with access
   b. Whether it should have known its actions were likely to interfere with access
   c. Whether it knew or should have known its actions were likely to interfere with access
   d. Whether its interference was malicious
3. Which of the following information blocking exceptions involves not fulfilling requests to access, exchange, or use EHI?
   a. Licensing
   b. Infeasibility
   c. Fees
   d. Content and manner
4. APIs (application programming interfaces) allow different software applications to:
   a. protect PHI against hackers
   b. use the same coded language

c. assess user fees
d. easily interact with one another

5. The Information Blocking Rule:
   a. focuses on appropriate information-sharing via the patient portal
   b. was created in response to patient portal practices that inhibit information-sharing
   c. does not mention patient portals
   d. specifically addresses the security of patient portals

Other Legal Considerations

Prior to HIPAA, many privacy protection regulations were handled at the state level. There were inconsistencies in these regulations among states if any existed at all. The HIPAA Privacy Rule provides a primary source of regulations for the privacy of protected health information. Other factors must be considered as well.

Preemption of State Law

Preemption of federal law is foundational to the U.S. Constitution and provides that, where there is conflicting state law, federal law will take precedence. This principle of course applies to HIPAA, which is federal law. There are exceptions, including when state law is more stringent. This means that state law provides individuals with greater privacy protections or greater rights regarding their PHI(45 CFR 160.202). For example, a state law that requires psychologists to obtain authorization for disclosure of information from new patients would prevail because it provides more protection for the PHI than HIPAA (which does not require an authorization) does.

However, there are numerous situations where state laws continue to take precedence even if they are less stringent than HIPAA because they serve recognized public policy purposes such as preventing healthcare fraud and abuse; regulating insurance and health plans; reporting on healthcare delivery and costs; serving compelling public health, safety and welfare needs; regulating controlled substances; or reporting disease or injury, child abuse, birth or death, or conducting public health surveillance (45 CFF 160.203).

E-Discovery

Traditionally, patient medical records documented much-needed information to support the medical care provided to the patient. Much more information is captured with electronic systems about the record itself, not just the healthcare data related to the care provided. This type of information is referred to as **metadata** and is the descriptive data that explains other data. In the context of the electronic health record (EHR), metadata can include the dates for the day the data were created, sent, received, or last accessed or modified. **E-discovery** refers to the Federal Rules of Civil Procedure and Uniform Rules Relating to Discovery of Electronically Stored Information. These rules address electronic information that is not usually considered a part of the legal EHR of a facility and is subject to motion for compulsory discovery. Audit trails, the source code of the program, metadata, and other electronic information are all electronically stored information and subject to discovery (AHIMA 2017).

Preparation for response to requests is paramount. Requests are increasing for health records along with the associated metadata. Preparation for these kinds of requests includes becoming as familiar with the EHR system as possible to identify when documented EHR events actually occurred, what is and is not accessible, and policies and procedures related to the creation, utilization, maintenance, retention, and destruction of

the information (Kirkendall et al. 2019). Revision of the information management plan, including the schedules for storage, retention, and destruction to address items specific to e-discovery, should be done. A number of policies and procedures will need to be updated or developed to support the information management plan. A legal hold policy, procedures, and employee training are important aspects of preparing for the eventual receipt of a request for the record and related metadata. A **legal hold** is "a communication issued because of current or anticipated litigation, audit, government investigation, or other such matters that suspend the normal disposition or processing of records" (AHIMA 2017). A legal hold for a paper health record can be simply locking it in a drawer to be preserved from alteration, changes, and such that are of concern when the record might be requested for legal purposes. With the EHR, there could be multiple systems, accessible to many authorized users, plus the metadata and other information that can be maintained in the EHR system, so there needs to be a means of freezing or preserving the record from potential alterations or destruction of the health record itself, plus the metadata. This process will involve communication with IT with a hold, freeze, preservation, or similar kind of notice, and they will then complete the steps necessary to preserve the record at a specific point in time.

Genetic Information

Genetic testing has become almost commonplace in healthcare. However, genetic information is not always understood, even by healthcare professionals. Because of the complexities of the information, as well as the potential consequences, the Genetic Information Nondiscrimination Act (GINA) was passed in 2008. GINA prohibits discrimination based on genetic information by health insurers related to eligibility, coverage, underwriting, or premium-setting decisions (Brodnik et al. 2023). Employers, employment agencies, labor organizations, and others are also prohibited from using genetic information for making employment decisions (Brodnik et al. 2023). This law has now been included in the 2013 HIPAA Omnibus amendments related to GINA to further define the terms health plan, underwriting purposes, and genetic information. Continued diligence in protecting the health information of consumers' genetic testing technologies is required as the uses for genetic information continue to evolve.

Case Study: Litigation and Metadata

A malpractice lawsuit was filed against a hospital in 2002, questioning the competency of a surgeon for performing the surgery that was done, which left the patient with a severe permanent complication. In reviewing the records obtained through e-discovery, issues in the anesthesia record were noted that changed the focus of the lawsuit to the anesthesiologist rather than the surgeon. There were 90 minutes where no vital signs were recorded, and the anesthesiologist documented his presence for the entirety of the surgery (seven hours) just minutes after the beginning of the surgery. This information was found due to the presence of time stamps embedded in the system. Although the anesthesiologist was present for the total time of the surgery, he recorded his surgery notes for the whole time at the beginning of the surgery, instead of as the events occurred. The complication suffered by the patient was not likely to have been caused or affected by the shortcomings identified in the data. Because it was known that the metadata could be used to discredit the anesthesiologist, the hospital settled the case with the patient before it was heard in court (Dimick 2007).

This case illustrates how having the metadata available for consideration in litigation can change the focus of a whole case. The original claim of malpractice against the hospital for incompetence by the surgeon was going to be difficult to prove due to the diminished state of the patient before surgery followed by a lengthy procedure. It was found that there were issues of documentation during the surgical procedure that were discovered with review of the metadata. This changed the focus of the case from the surgeon to the anesthesiologist, and even though his documentation issues did not cause the patient's complication, this information was enough to discredit his competency and the case was settled by the hospital. The care provided was good, but the issue with the timing of the documentation cast doubt on the quality.

Administrative Requirements for Covered Entities

There are several administrative requirements that must be met to comply with HIPAA and the HITECH Act. These requirements include policies and procedures to comply with the HIPAA Privacy Rule, notices of privacy practice, data safeguards, mitigation protocol, designation of a privacy officer, workforce training, and accounting of disclosures.

Notice of Privacy Practices

The **Notice of Privacy Practices (NPP)** is an essential statement by a healthcare organization that informs individuals of the uses and disclosures of patient-identifiable health information that may be made by the organization, as well as the individual's rights and the organization's legal duties with respect to that information (AHIMA 2017). The NPP is an important tool to help patients understand their rights. Many model NPPs are available at the Health IT website for facilities to refer to when preparing their own, but care should be taken to assure the final NPP is specific to that organization (HealthIT 2016).

Policies and Procedures

Policies and procedures are required to ensure compliance with the Privacy Rule. An ongoing review of privacy policies and procedures, their implementation, and consistency with changes in the privacy and security regulations must be completed. A CE's Notice of Privacy Practices must reflect material regulatory changes (Brodnik et al. 2023).

Training

Workforce members must be trained on HIPAA policies and procedures so they can complete their job duties appropriately and lawfully. Because the Privacy Rule emphasizes documentation, all training should be ongoing and documented for each employee (Brodnik et al. 2023).

Privacy Officer

Per 45 CFR 164.530(a)(1), a CE must designate a person responsible to develop and implement privacy policies and procedures. This could be a full-time position or added to another's job duties. There must also be a contact person or office responsible for receiving complaints (Brodnik et al. 2023).

Safeguards

CEs are required to ensure appropriate administrative, technical, and physical safeguards to protect the privacy of PHI (45 CFR 164.530(c)). Reasonable safeguards for paper records include isolating and locking file cabinets or record storage rooms. Safeguards for EHRs include passwords, firewalls, and workstation cubicles (Brodnik et al. 2023).

Mitigation

To the extent that it is realistic, CEs must mitigate known harmful effects due to the use or disclosure of PHI (45 CFR 164.530). **Mitigation** is the process of moderating the severity and potential harmful effects to an organization from wrongful use and disclosure. The protocol establishes a plan for potential courses of action, which could include such steps as extending an apology, disciplining the responsible party, updating procedures, and gestures of good public relations (Brodnik et al. 2023).

Retaliation and Waiver

CEs are expressly prohibited from **retaliation and waiver**, which suppress an individual's ability to complain about alleged HIPAA Privacy Rule violations (Brodnik et al. 2023). This means that CEs may not retaliate against anyone who exercises his or her rights under the Privacy Rule, assists in an HHS or other investigation, or opposes an act the person believes is a violation of the Privacy Rule (45 CFR 164.530(g)). Further, individuals cannot be mandated to waive their Privacy Rule rights in order to obtain treatment, payment, or enrollment/benefits eligibility (45 CFR 164.530(h)).

Documentation and Record Retention

Documentation related to the Privacy Rule must be retained six years from the date the document was created or the last effective date of the document, whichever is the most recent (45 CFR 164.530(j)). Relevant documents that must be maintained include policies and procedures, the NPP, complaints and their dispositions, and other documentation required by the Privacy Rule (Brodnik et al. 2023).

Enforcement and Penalties for Non-Compliance

The OCR enforces the Privacy Rule through the final HIPAA **Enforcement Rule**, which addresses compliance, investigations, civil monetary penalties, and procedures related to hearings (HHS 2020). **Resolution agreements** require CEs or BAs to perform obligations per settlements that often include payments and OCR monitoring. Settlements include **civil monetary penalties (CMPs)**, which are fines resulting from HIPAA violations (Brodnik et al. 2023).

Civil monetary penalties have levels, or tiers, that are based on intent and neglect. The nature and extent of both the violation and the harm are used to determine the amount of the fine. Fines are adjusted annually for inflation:

- Tier 1: \$100–\$50,000 per violation for unknowing violations (would not have known a violation was committed, even with reasonable diligence), capped at \$25,000/year
- Tier 2: \$1,000–\$50,000 per violation if due to reasonable cause (CE or BA knew or would have known with reasonable diligence), but not willful neglect, capped at \$100,000/year
- Tier 3: \$10,000–\$50,000 per violation if due to willful neglect and corrected within 30 days of discovery, capped at \$250,000/year
- Tier 4: \$50,000 per violation if due to willful neglect and not corrected as required, capped at \$1.5 million/year (Alder 2022)

Breaches

Prior to HITECH, wrongful uses and disclosures did not have to be reported to the affected individuals or to the government. However, breach determination and breach notification are now hallmarks of HIPAA (Brodnik et al. 2023).

Definition of Breach

Per HIPAA, a **breach** is an unauthorized "acquisition, access, use or disclosure of PHI, […] which compromises the security or privacy of such information" (45 CFR 164.402). An impermissible use or disclosure of PHI is presumed to be a breach unless there is a low probability that the PHI has been compromised. Low probability must be decided through an organization's breach determination process.

"Low probability" is determined by conducting a risk assessment that considers these factors:

1. The nature and extent of the PHI involved, including the types of identifiers and the likelihood of re-identification.
2. The unauthorized person who used the PHI or to whom the disclosure was made.
3. Whether the PHI was actually acquired or viewed.
4. The extent to which the risk to the PHI has been mitigated. (45 CFR 164.402)

Breaches apply only to unsecured PHI, which technology has not made unusable, unreadable, or indecipherable to unauthorized persons (45 CFR 164.402). As of now, encryption is a technology that safeguards PHI against breaches because it codes or scrambles data to make them indecipherable (Brodnik et al. 2023).

Exceptions to the breach definition include disclosures to unauthorized recipients if they would not reasonably be able to retain the disclosed information; situations where a workforce member or individual, acting under the CE's or BA's authority, unintentionally acquires, accesses, or uses the PHI if it was in good faith, within the scope of authority, and could not be further disclosed or used in an impermissible manner; or there was inadvertent disclosure by an individual at a CE or BA to another authorized person at a CE or BA, and the information is not further disclosed or used in an impermissible manner (45 CFR 164.402).

When all of the above conditions have been met and there are no exceptions, a breach must be reported and affected individuals must be notified.

Breach Notification

Breaches are "discovered" when they are first known, or they should reasonably should have been known. Individuals whose information has been breached must be notified without unreasonable delay (no more than 60 days), by first-class mail and by a faster method (such as telephone) if there is the potential for "imminent misuse." Where more than nine individuals are affected and written notice is unsuccessful, web postings or the media are recommended. When 500 individuals are affected, media outlets must be notified and the Secretary of HHS must be notified immediately. OCR provides an online breach reporting system that must be used to report all breaches (ie, one or more individuals). With the exception of breaches involving 500+ individuals, where immediate notification to OCR is required, all breaches within a calendar year must be entered into the online system within 60 days of the calendar year after which the breaches occurred (Brodnik et al. 2023).

Individuals whose PHI has been breached must be provided with the following information:

- A description of what occurred (including date of breach and date that breach was discovered)
- The types of unsecured PHI that were involved (such as name, social security number, date of birth, home address, and account number)
- Steps that the individual may take to protect himself or herself
- What the entity is doing to investigate, mitigate, and prevent future occurrences
- Contact information for the individual to ask questions and receive updates (45 CFR 164.404)

page 291

HIPAA CEs and BAs, which may include personal health record (PHR) vendors, are subject to HHS-issued breach notification regulations. Non-CEs and non-BAs (including non-BA PHR vendors, third-party service providers of PHR vendors, and other non-HIPAA CEs or BAs affiliated with PHR vendors including makers of health apps) are subject to companion breach notification regulations issued by the Federal Trade Commission (FTC) (FTC 2022).

Notification Audit Program

National standards for the privacy and security of PHI were established with HIPAA. Breach notification requirements were added by the HITECH Act to provide increased transparency for those whose information could be at risk. Periodic audits related to compliance by the BAs and CEs became a required activity to be conducted by the OCR. page 292 Phase 1 of the program involved the implementation of a pilot program for the assessment of implemented controls and processes, followed by a comprehensive evaluation review of the pilot program. Phase 2 of the program went into effect in 2016 for the next phase of audits to analyze implemented policies and procedures of designated standards and specifications of CEs and BAs.

Table 10.3 . Audit protocol for HIPAA Section 164.502(i)

Audit type	Section	Key activity	Established performance criteria	Audit inquiry

Privacy	§ 164.502(i)	Uses and disclosures consistent with notice	§ 164.502(i) Standard: Uses and disclosures consistent with notice: A covered entity that is required by § 164.502 to have a notice may not use or disclose protected health information in a manner inconsistent with such notice. A covered entity that is required by § 164.502(b)(1)(iii) to include a specific statement in its notice if it intends to engage in an activity listed in § 164.502(b)(1)(iii)(A)-(C) may not use or disclose protected health information for such activities, unless the required statement is included in the notice.	Are uses and disclosures made by the covered entity consistent with its notice of privacy practices? Inquire of management whether uses and disclosures of PHI are consistent with the entity's notice of privacy practices. Obtain and review policies and procedures regarding uses and disclosures. Evaluate whether the uses and disclosures of PHI are consistent with the entity's notice of privacy practices.

Source: HHS (2020)

Breach Audit Protocol

The reviews for Phase 2 will be based on an updated audit protocol that reflects the 2013 Omnibus Final Rule. An example of the audit protocol for one section to be audited is shown in Table 10.3.

The audit program is important for meeting compliance requirements and provides means to assess compliance activities carried out by the CEs and BAs. The audits also provide information on the processes of how they are meeting compliance, discover issues that might not have been recognized through the investigation of breach complaints, and aid in the development of best practices, all of which will improve the CEs and BAs with compliancy in the future.

Chapter 10 Review Exercises

1. State laws can pre-empt HIPAA:
   a. If the state law was written first, before HIPAA
   b. At the discretion of the facility and written in hospital policy
   c. If they are more stringent than HIPAA
   d. Never; HIPAA always prevails
2. An accounting of disclosures must include:
   a. Identifiable health information used in the facility's internal operations
   b. Incidental disclosures
   c. Disclosures pursuant to an authorization
   d. Disclosures for public health purposes

3. Which of the following will not allow state law to prevail over HIPAA?
    a. Necessary to prevent fraud and abuse related to payments
    b. Purpose is to regulate manufacture, registration, and distribution of controlled substances
    c. For the reporting of diseases, injury, child abuse, birth, or death
    d. When state laws are contrary to the Privacy Rule
4. Data safeguards, mitigation protocol, and designation of a privacy offer are:
    a. Physical safeguards
    b. Administrative requirements
    c. Policies and procedures
    d. At the discretion of the facility
5. E-Discovery rules address:
    a. Searching the patient health record electronically
    b. Requesting that the requested record be sent electronically
    c. The release of the patient health information of that maintained electronically
    d. Includes requests for electronically stored information to include metadata

6. Genetic information:
    a. Can be used to make health insurance coverage decisions
    b. can be used to make employment consideration
    c. is protected by GINA
    d. may be taken into consideration to establish health insurance rates
7. The primary focus of phase 2 of the breach audit protocol is:
    a. Compliance
    b. Monitoring notices of privacy practices
    c. Assessing controls and processes
    d. Reviewing policies and procedures

Discussion

8. Briefly explain the intent of the breach notification audit program.
9. Explain the differences between the use and disclosure of PHI.
10. Discuss the intent of the HIPAA Privacy Rule.
11. Discuss how ePHI has evolved with the advances of technology
12. The EHI definition excludes psychotherapy notes and information compiled in reasonable anticipation of or for use in a civil, criminal, or administrative action or proceeding. Discuss why these records are excluded as EHI.
13. Select one of the information blocking exceptions and explain why the authors of the Information Blocking Rule included it as an exception.

REFERENCES

42 CFR 164.512: Privacy Law. 2013.

42 CFR 493.3(a)(2): Laboratory requirements. 2013.

45 CFR 160.103: Definitions. 2013.

45 CFR 160.202: Definitions. 2013.

45 CFR 160.203: General rule and exceptions. 2013.

45 CFR 162.103. Definitions. 2015.

45 CFR 164.308(b)(1): Administrative safeguards. 2013.

45 CFR 164.501: Definitions. 2013.
45 CFR 164.502: Uses and disclosures of protected health information: General rules. 2013.
45 CFR 164.502(b): Uses and disclosures of protected health information: General rules. 2013.
45 CFR 164.504: Uses and disclosures: Organizational requirements. 2013.
45 CFR 164.506: Uses and disclosures for treatment, payment and operations: General rules. 2013.

45 CFR 164.508: Uses and disclosures for which an authorization is required. 2013.
45 CFR 164.508(a)(2): Uses and disclosures for which an authorization is required. 2013.
45 CFR 164.514(b): Implementation specifications. 2013.
45 CFR 164.524: Access of individuals to protected health information. 2013.
45 CFR 164.528(a)(1): Accounting of disclosures of protected health information. 2013.
45 CFR 164.528(b): Accounting of disclosures of protected health information. 2013.
45 CFR 164.528(d): Accounting of disclosures of protected health information. 2013.
45 CFR 164.530(a)(1): Administrative requirements. 2013.
45 CFR 164.530(c): Administrative requirements. 2013.
45 CFR 164.530(f): Standard: Mitigation. 2013.
42 CFR 171.102: Definitions. 2020.
45 CFR 171.103: Information Blocking. 2020.
AHIMA. 2015. ***Consumer Health Information Bill of Rights***. http://bok.ahima.org/PdfView?oid=107674.
AHIMA. 2017 ***Pocket Glossary for Health Information Management and Technology***, 5th ed. Chicago: AHIMA Press.
Belfort, R. 2010. "***HITECH Raisesthe Stakes on HIPAA Compliance."*** https://www.manatt.com/insights/newsletters/health-highlights/_healthcare-law_07-13-10.
Brodnik, M., L. Rinehart-Thompson, and R. Reynolds. 2023. Fundamentals of Law for Health Informatics and Information Management, 4th ed. Chicago: AHIMA.
Department of Health and Human Services. 2011. HIPAA Privacy Rule accounting of disclosures under the Health Information Technology for Economic and Clinical Health Act. ***Federal Register*** 76(104):31426–31449.
Department of Health and Human Services (HHS). 2013. "***Uses and disclosures for treatment, payment, and health care operations.***" https://www.hhs.gov/hipaa/for-professionals/privacy/guidance/disclosures-treatment-payment-health-care-operations/index.html .
Department of Health and Human Services (HHS). 2016. ***Individuals' Right under HIPAA to Access their Health Information.*** 45 CFR § 164.524. Health Information Privacy.
Department of Health and Human Services. 2020. "***HIPAA Privacy, Security, and Breach Notification Audit Program.***" https://www.hhs.gov/hipaa/for-professionals/compliance-enforcement/audit/index.html.
https://www.hhs.gov/hipaa/for-professionals/privacy/guidance/disclosures-treatment-payment-health-care-operations/index.html .
Dimick, C. 2007. e-Discovery: Preparing for the coming rise in electronic discovery requests. ***Journal of AHIMA*** 78:24–29.

Dworkowitz, A. 2022. Provider Obligations for Patient Portals Under The 21st Century Cures Act. ***Health Affairs***. www.healthaffairs.org.

Food and Drug Administration (FDA). 2020. 21st Century Cures Act. https://www.fda.gov/regulatory-information/selected-amendments-fdc-act/21st-century-cures-act#.

Glaser, John and E. Gardner. 2022. Standardized APIs Could Finally Make It Easy to Exchange Health Records. ***Harvard Business Review***. https://hbr.org/2022/03/standardized-apis-could-finally-make-it-easy-to-exchange-health-records.

Health Insurance Portability and Accountability Act of 1996. Public Law 104-191.

Health IT. 2016. "***Model notices of privacy practices.***" https://www.healthit.gov/providers-professionals/model-notices-privacy-practices.

HIPAA Administrative Simplification Statute and Rules. 2021. "HIPAA for professionals." http://www.hhs.gov/ocr/privacy/hipaa/administrative/index.html.

HIPAA Privacy Rule, Part 160, 2022. https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-C/part-160?toc=1 .

Kirkendall, E. S., Y. Ni, T. Lingren, M. Leonard, E. S. Hall, and K. Melton. 2019. Data challenges with real-time safety event detection and clinical decision support. Journal of Medical Internet Research 21(5):e13047. doi:10.2196/13047.

Office for Civil Rights, Department of Health and Human Services (OCR). 2013a. Modifications to the HIPAA privacy, security, enforcement, and breach notification rules under the Health Information Technology for Economic and Clinical Health Act and the Genetic Information Nondiscrimination Act; other modifications to the HIPAA rules. ***Federal Register*** 78(17):5566–5702.

Office for Civil Rights, Department of Health and Human Services (OCR). 2013b. "***Summary of the HIPAA Privacy Rule.***" https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html.

Office for Civil Rights, Department of Health and Human Services. (OCR). 2013. ***Summary of the HIPAA Privacy Rule***. https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html.

Office for Civil Rights, Department of Health and Human Services. (OCR). 2017. ***Covered Entities and Business Associates.*** https://www.hhs.gov/hipaa/for-professionals/covered-entities/index.html.

Office for Civil Rights, Department of Health and Human Services. (OCR). 2019. ***Business Associates.*** https://www.hhs.gov/hipaa/for-professionals/privacy/guidance/business-associates/index.html.

Office for Civil Rights, Department of Health and Human Services. (OCR). 2022a. "***Guidance regarding methods for de-identification of protected health information in accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule.***" https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html.

Office for Civil Rights, Department of Health and Human Services (OCR). 2022b. "Eleven Enforcement Actions Uphold Patients' Rights Under HIPAA." https://www.hhs.gov/about/news/2022/07/15/eleven-enforcement-actions-uphold-patients-rights-under-hipaa.html .

Office of the National Coordinator for Health Information Technology (ONC). 2022. Information Blocking. https://www.healthit.gov/topic/information-blocking.

Rockwern, B., D. Johnson, L. S. Sulmasy, and Medical Informatics Committee and Ethics, Professionalism and Human Rights Committee of the American College of Physicians. 2021. Health information privacy, protection, and use in the expanding digital health ecosystem: A position paper of the American College of Physicians. Annals of Internal Medicine 174(7):994–998. doi:10.7326/M20-7639.

Sullivan, J., and S. Hartsfield. 2020. ***HIPAA: A Practical Guide to the Privacy and Security of Health Data***, 2nd ed. Chicago: American Bar Association.

Tripathi, M. and S. Posnack. 2021. A New Day for Interoperability – The Information Blocking Regulations Start Now. HealthITbuzz. https://www.healthit.gov/buzz-blog/information-blocking/a-new-day-for-interoperability-the-information-blocking-regulations-start-now .

ADDITIONAL RESOURCES

45 CFR 164.508(b)(2): Uses and disclosures for which an authorization is required. 2013.

45 CFR 164.510: Uses and disclosures requiring an opportunity for the individual to agree or to object. 2013.

45 CFR 164.514: Other requirements relating to uses and disclosures of protected health information. 2013.

45 CFR 164.514(b)(2): Other requirements relating to uses and disclosures of protected health information. 2013.

Breach Notification Rule. 2013. https://www.hhs.gov/hipaa/for-professionals/breach-notification/index.html.

Brocato, L., S. Emery, and J. McDavid. 2011. Keeping compliant: Managing rising risk in physician practices. ***Journal of AHIMA*** 82(11):32–35.

Centers for Medicare and Medicaid Services (CMS), Office of E-Health Standards and Services (OESS). 2008. ***HIPAA Compliance Review Analysis and Summary of Results***. http://www.hhs.gov/ocr/privacy/hipaa/enforcement/cmscompliancerev08.pdf.

Department of Health and Human Services, Office for Civil Rights (OCR). 2003. "***Protected health information.***" http://www.hhs.gov/ocr/privacy/hipaa/understanding/training/udmn.pdf.

Guidance on Risk Analysis Requirements under the HIPAA Security Rule. 2010. http://www.hhs.gov/ocr/privacy/hipaa/administrative/securityrule/rafinalguidancepdf.pdf.

Heubusch, K. 2011. Access report: OCR tries subtraction through addition in accounting of disclosure rule. ***Journal of AHIMA*** 82(7):38–39.

HIPAA Compliance Review Analysis and Summary of Results. 2008. http://www.hhs.gov/ocr/privacy/hipaa/enforcement/cmscompliancerev08.pdf.

Ouellette, P. 2014. HIPAA Privacy Rule: Permitted PHI uses and disclosures. ***Health IT Security***. http://healthitsecurity.com/news/hipaa-privacy-rule-permitted-phi-uses-and-disclosures.

Recent Changes to HIPAA—The HITECH Act. 2022. https://www.hipaajournal.com/new-hipaa-regulations/.

Chapter

11

Security for Health Information

By Leming Zhou, PhD

Learning Objectives

- Interpret the HIPAA Security Rule
- Apply security risk analysis activities
- Analyze HIPAA security risk safeguards
- Articulate the importance of confidentiality, integrity, and availability in regard to the HIPAA Security Rule
- Relate HIPAA Security Rule penalties and enforcement process
- Explain how HITECH modifies the HIPAA Security Rule
- Summarize medical identity theft and its potential impacts
- Develop plans for disaster preparedness

KEY TERMS

Addressable implementation specifications
Administrative safeguards
Availability
Breach
Break the glass
Business continuity planning (BCP)
Business impact analysis
Civil penalties
Cold site
Confidentiality
Criminal penalties
Disaster recovery planning (DRP)

Encryption
Health Information Technology for Economic and Clinical Health (HITECH) Act
Health Insurance Portability and Accountability Act (HIPAA)
HIPAA Enforcement Rule
Hot site
Implementation specifications
Integrity
Medical identity theft
Office for Civil Rights (OCR)
Physical safeguards
Recovery point objective (RPO)
Recovery time objective (RTO)
Red flags
Required implementation specifications
Risk
Sanctions policy
Security incident
Security risk analysis
Standards
Technical safeguards
Threat
Vulnerability
Warm site

page 298

The Health Insurance Portability and Accountability Act (HIPAA) of 1996 requires that patient identifiable health information is protected and that protected health information (PHI) is available as needed for the delivery of care. There are multiple standards in HIPAA, such as the Privacy Rule, the Security Rule, Electronic Transactions and Code Set Rule, and national identifier requirements for employers, providers, and health plans (HIPAA 1996). Building on HIPAA, the **Health Information Technology for Economic and Clinical Health (HITECH) Act** outlines breach notification requirements and starts to lay the groundwork for additional protections as healthcare continues to move toward an integrated health information exchange environment.

A **breach** is defined under the 2013 Final Rule as "the acquisition, access, use, or disclosure of protected health information in a manner not permitted which compromises the security or privacy of the protected health information" (45 CFR 164.402). From March 2019 to March 2022, there were 1,035 breaches of electronic protected health information (ePHI) documented in an investigation conducted by the Office for Civil Rights (OCR) in the US Department of Health and Human Service (HHS). Each of these breaches affected at least 500 individuals. The largest one affected more than 11.5 million patients. The majority of these breachers were due to hacking or IT incidents (795 in total, or 76.8 percent; see figure 11.1; OCR 2022a). Although some hacking is highly sophisticated, most hacking could be prevented or mitigated if HIPAA covered entities (CEs) and business associates (BAs) fully implemented HIPAA Security Rule requirements to handle the most common types of attacks, such as phishing emails, exploitation of known vulnerabilities, and weak authentication protocols (OCR 2022b).

page 299

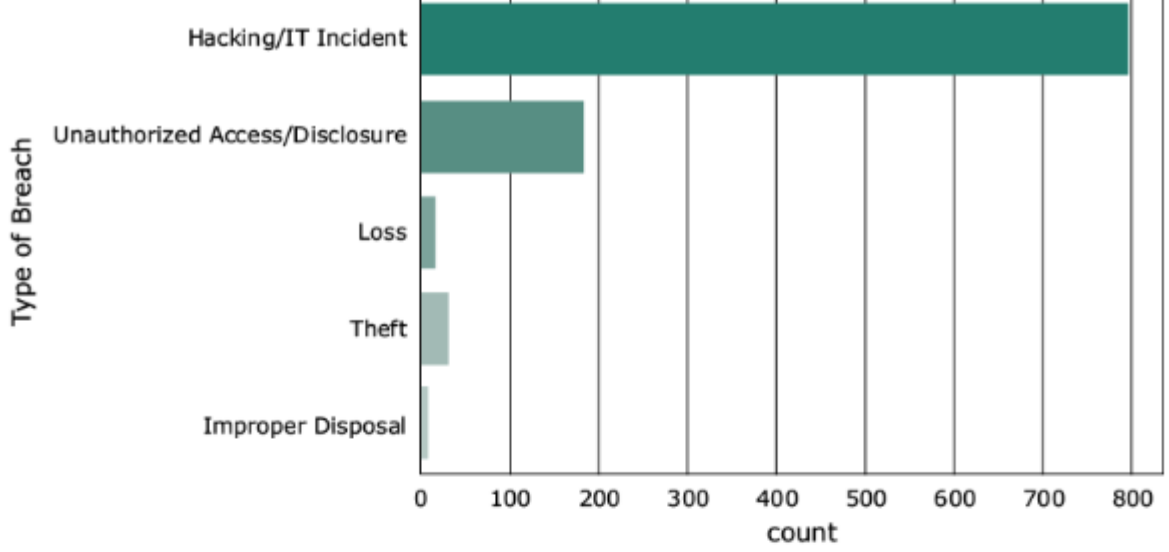


Figure 11.1. The type of breaches under investigation from 2019 to 2022

▶ Expand/collapse Extended Description

The Security Rule covers all ePHI created, received, maintained, or transmitted by an organization (OCR 2010). The **standards** and **implementation specifications** of the Security Rule provide a baseline for protecting ePHI. The Security Rule provides a foundation for helping CEs and BAs ensure the confidentiality, integrity, and availability of their ePHI.

This chapter introduces specific guidelines to assure compliance with the security requirements for electronic health information. The enforcement of the guidelines is covered, as well as guidelines for handling breaches if they occur. Consequently, protection against medical identity theft and disasters are two considerations for a facility to ensure that they will be able to have the correct information available in a timely manner to support patient care.

Security Rule

In the Security Rule, the required standards include required and addressable implementation specifications intended to provide CEs and BAs with flexible, scalable, technology-neutral solutions and alternatives for complying with the standards.

Required implementation specifications must be implemented in conformance with the regulation. The term addressable implementation specification does not imply that the specification is optional. **Addressable implementation specifications** should be implemented unless an organization determines that the specification is not reasonable or appropriate for the specific organization. If this is the case, the organization must document why it is not reasonable or appropriate and adopt an equivalent measure if it is reasonable and appropriate to do so (45 CFR 164.306(d)(3)).

In addition to the extensive language in the Rule intended to secure ePHI, a **security incident** is defined as "the attempted or successful unauthorized access, use, disclosure, modification, or destruction of information or interference with system operations in an information system" (45 CFR 164.304). This definition is important as it serves as the method by which to evaluate reported Security Rule violations.

The foundational basis of a CE's and BA's information security program rests on its identification and management of potential risks to ePHI. The Security Rule outlines a

broad set of regulations that include administrative, technical, and physical safeguards intended to ensure confidentiality, integrity, and availability of ePHI. Even in the best of situations, risk can turn into reality. For this reason, the Security Enforcement Rule outlines penalties and enforcement for violations of the rule.

Security Risk Analysis

The HIPAA Security Rule requires full evaluation of the methods, operational practices, and policies used by the CE to secure ePHI (45 CFR 164.302–318). CEs' and BAs' obligations regarding risk analysis are outlined in 45 CFR 164.308(a)(1)(ii)(A). The **security risk analysis** process provides CEs and BAs with the structural framework on which to build their HIPAA security plan. CEs and BAs are required to conduct a security risk analysis "to evaluate risks and vulnerabilities in their environment and to implement reasonable and appropriate measures to protect against reasonably anticipated threats or hazards to the security or integrity of the e-PHI" (OCR 2010).

The value of a risk analysis stems from its uniqueness to the specific organization in which it is conducted. Every organization is different, and every risk analysis should reflect the unique and complex interrelationships between a multitude of systems, processes, and policies that, in combination, result in that specific organization's HIPAA security plan. For this very reason, the Security Rule is not prescriptive in requiring any specific method or approach for conducting the risk assessment. The Risk Analysis Final Guidance Document spells out some essential questions organizations should consider during the risk analysis process (OCR 2010). These questions arise from the National Institute of Standards and Technology (NIST) Special Publication (SP) 800-66 and include the following:

- Have you identified the ePHI within your organization? This includes ePHI that you create, receive, maintain, or transmit.
- What are the external sources of ePHI? For example, do vendors or consultants create, receive, maintain, or transmit ePHI?
- What are the human, natural, and environmental threats to information systems that contain ePHI? (NIST SP 800-66 2008)

The ultimate goal of the risk analysis process is to guide organizations in the decisions made and actions taken to comply with the Security Rule's standards and addressable or required implementation specifications. Although necessary to put the risk analysis discussion in context, the concepts of **vulnerability,** threats, and risks are not expressly defined in the Security Rule. The Risk Analysis Guidance Document again refers to the NIST SP 800-66 document to frame the risk analysis discussion.

Vulnerabilities

Vulnerability is defined as an inherent weakness or absence of a safeguard that could be exploited by a threat (AHIMA 2017, 152). Vulnerabilities are grouped into either technical or nontechnical categories. Technical vulnerabilities reflect inappropriate information systems that potentially provide the means for assault, harm, or unauthorized corruption. Examples of technical vulnerabilities can be software program errors in the computer systems, computer hardware failure, and legacy computer systems without necessary updates. Nontechnical vulnerabilities are demonstrated by such things as policy and procedure weaknesses. Examples of nontechnical vulnerabilities are lack of regular training in security and privacy, allowing many people to have privileged access to the electronic health record (EHR) systems, and no proper procedures for patient information release or employee termination.

Hackers can gain access to ePHI by exploiting known vulnerabilities. A known vulnerability is a vulnerability whose existence is publicly known. NIST (2022) maintains the National Vulnerability Database, which has an extensive list of known vulnerabilities and their information. Typically, after a vulnerability is made publicly known, vendors would create software patches or release a new version of the software system to mitigate the vulnerability. If a CE or BA chooses not to install the patches or update their system, however, the vulnerability would remain and a hacker can penetrate the CE's or BA's network and gain access to ePHI.

Threats

A **threat** is the potential for exploitation of a vulnerability or potential danger to a computer, network, or data (AHIMA 2017, 145). Threats are grouped into three categories: natural, human, and environmental. Natural threats are often weather related and include tornadoes, hurricanes, earthquakes, floods, and the like. Human threats are broad in nature and can be intentional or unintentional. Access to information by unauthorized individuals is a good example of a human threat because this can occur through intentional actions of the user, or can result from an accidental key entry error that exposes information to an unauthorized user. Environmental threats center on power failure or chemical or other environmental agents with the potential to damage electronic data. With the improvement of information technologies, especially the wide adoption of the cloud-based services in healthcare, the impact from natural and environmental threats is not as severe as it was in the past, but the human threats are still as serious as ever.

Risks

Risk is the probability of incurring injury or loss, the probable amount of loss foreseen by an insurer in issuing a contract, or a formal insurance term denoting liability to compensate individuals for injuries sustained in a healthcare facility (AHIMA 2017, 131).

Securing ePHI can be very costly and time consuming. It might not be possible to eliminate all threats, so in those cases, organizations must determine the extent of effort needed to reduce and mitigate the remaining risk. When there is a low probability of the occurrence of a risk and the impact of the potential risk is low, the risk would then be considered a low-level risk. Conversely, if the probability of the risk is high, along with a high impact due to the risk, the risk would then be considered critical.

Risk Analysis Methods

The Security Rule intentionally leaves the methods for conducting the required risk analysis to the discretion of the entity. Regardless of the methods selected for conducting and documenting the risk analysis, the Security Rule does mandate several elements that must be included in the analysis:

- Define scope of the risk analysis
- Identify means of data collection
- Identify and document potential threats and vulnerabilities
- Assess current security measures
- Determine the likelihood of threat occurrence
- Determine the potential impact of threat occurrence
- Determine the level of risk
- Finalize documentation
- Perform periodic review and updates to the risk assessment (45 CFR 164.308(a)(1))

Even though these elements must be included, they are to be considered a guide and a communication of the expectations. A healthcare facility should first consider its own characteristics and environment before proceeding with a risk analysis. Once this self-assessment has been completed, the healthcare facility can then determine the best way to conduct a risk analysis specific to itself and its situation while maintaining compliance with the mandated elements.

HIPAA Security Rule Safeguards

The Security Rule outlines a variety of safeguards necessary to ensure appropriate management and protection of ePHI. Working in concert, administrative, physical, and technical safeguards can produce the intended outcome of preserving the confidentiality, integrity, and availability of ePHI. Although the intent and approach of each of these forms of safeguards vary greatly from one another, each form of safeguard is necessary.

Administrative Safeguard Standards

Administrative safeguards are administrative actions, policies, and procedures to manage the selection, development, implementation, and maintenance of security measures to protect ePHI and manage the conduct of the CE's or BA's workforce in relation to the protection of that information (45 CFR 164.304). The specific administrative safeguard standards are discussed next.

Security Management Process Standard

The administrative safeguards comprise over half of all the safeguards included in the Security Rule. Many believe the administrative safeguards are the most difficult to address as they rely heavily on human involvement to ensure compliance with the safeguards.

The first standard in the administrative safeguards section of the Security Rule is the security management process, which is found at 45 CFR 164.308(a)(1). This broad standard has four required implementation specifications that include the previously discussed risk analysis, a required risk management element, a required sanctions policy, and a required information systems activity review element. The risk management implementation specification picks up where risk analysis left off and outlines how the identified risks will be managed. Two critical factors in the risk management process include communication of security processes and leadership involvement with risk mitigation.

The **sanctions policy** must outline how cases of noncompliance will be addressed within the organization, knowing that even with a strong risk management plan, the potential for noncompliance (either intentional or nonintentional) still exists. Some of the components of the sanctions policy are a discussion of the significance of noncompliance, examples of noncompliance, and a sliding scale of discipline based on the severity of the act of noncompliance.

The implementation specification requires CEs to "implement procedures to regularly review records of information system activity, such as audit logs, access reports, and security incident tracking reports" (CMS 2007a). Today's information systems have the ability to track and record a multitude of user activities, but unless specific measures are enacted by the organization to review these tracking progeny, security violations could go unnoticed. This implementation specification advises CEs to consider not only the technical abilities of their systems to create audit logs and reports, but also to address the nontechnical use of these tools through policy and procedures.

Security Officer

Assigned security responsibility is the second administrative safeguard standard detailed in 45 CFR 164.308(a)(2). This standard requires CEs to "identify the security official who is responsible for the development and implementation of the policies and procedures required by this subpart [the Security Rule] for the entity" (45 CFR 164.308(a)(2)). CEs must establish who in their organization is responsible for compliance with the Security Rule requirements. This position is similar in scope and responsibility to that of the privacy officer required by the HIPAA Privacy Rule. CEs may (but are not required to) appoint the same individual to fill the role of both the privacy and the security officer.

Workforce Security Standard

Workforce security is the third administrative safeguard standard and has three associated addressable implementation specifications (45 CFR 164.308(a)(3)). This standard requires the organization to ensure that those with a legitimate need to access information are able to do so, at the same time ensuring that those workforce members who do not have a legitimate need for that information are prevented from gaining access.

- The first implementation specification for this standard is authorization and supervision, which speaks to the unique needs of an organization based on size, function, and structure, and spells out how each organization must determine each workforce member's level of access.
- The second implementation specification, workforce clearance procedures, states that CEs must "implement procedures to determine that the access of a workforce member to electronic protected health information is appropriate" (45 CFR 164.308(a)(3)).
- The third implementation specification in this section is termination procedures, which challenges CEs to implement appropriate methods to ensure workforce members' access privileges are removed when they terminate employment (willingly or unwillingly).

This implementation specification also outlines the need to ensure changes in employee scope and function be addressed by policy and procedures that facilitate timely revisions to access level as indicated by position changes.

The workforce of an organization poses a particular risk in regard to security due to the ability of employees to access PHI and the innate nature of people's desire to know. Organizations address this risk with new employee orientation, policies and procedures, employee handbooks, continuing education, and monitoring and follow-up. Although all of these are carried out in healthcare organizations, there continue to be instances of violations of the Security Rule.

The types of data breaches listed here are just a few examples of the kinds of breaches that are commonly made by employees working in a healthcare facility.

- An employee might have inappropriately accessed thousands of patients' health records as determined by an internal audit. There was no indication that the information was disclosed to a third party or used for malicious purposes, such as identity theft. The employee was terminated per hospital policy.
- Patient information was taken by a former employee when leaving employment from the facility. The affected patients were notified.
- A health system employee accessed patient records inconsistent with her job functions, using the information to file claims with other employees' health insurance companies. Affected individuals were notified and the employee was terminated.

- Fourteen employees of a clinic inappropriately accessed a well-known patient's records. All employees were investigated to determine their role and level of participation in accessing the information and all were disciplined, with some terminated from employment.
- A vendor's employee accessed patient information unrelated to his job duties.
- A hospital employee was found to have hospital documents with patient information in his home. The employee used the patient information to apply for credit cards.

These information breaches represent a few of the many ways that members of the healthcare workforce are a leading cause of PHI breaches. These workers have access to the EHR systems to meet the obligations of their job but that does not give them the right to access information on all patients and certainly not to use the PHI for personal gain. The workforce security safeguard is a crucial component for achieving compliance with the Security Rule.

Information Access Management Standard

The fourth administrative safeguard standard is the information access management standard (45 CFR 164.308(a)(4)). This standard works in tandem with 45 CFR 164.308(a)(3) to uphold the basic security tenet of restricting access to only those who need it to perform their jobs. However, this standard specifically calls out a required implementation specification for isolating healthcare clearinghouse functions to ensure that clearinghouses that are part of a larger organization have their own policies and procedures to segregate ePHI access from unauthorized access by the larger organization.

Two addressable implementation specifications for access authorization and access establishment and modification are also included in the information access management standard. In terms of access authorization, the previous workforce security standards established workforce member rights to access information; the access authorization implementation specification addresses the policies and procedures for granting access to ePHI. Parallels are also drawn between the workforce security implementation specifications and the authorization, and access establishment and modification implementation specification. This implementation specification is, "Implement policies and procedures that, based upon the entity's access authorization policies, establish, document, review, and modify a user's right of access to a workstation, transaction, program, or process" (45 CFR 164.308(a)(4)).

Security Awareness and Training Standard

The fifth administrative safeguard standard requires CEs to "implement a security awareness and training program for all members of its workforce (including management)" (45 CFR 164.308(a)(5)). Four addressable implementation specifications support this standard and provide ample flexibility for CEs to meet both the intent and spirit of this standard. This standard requires all existing workforce members to receive security training and periodic retraining as indicated by changing conditions and training for new staff. All training must be documented. For instance, the security training can be offered online, but the training system must have a user authentication mechanism so that the system can be certain that the person who receives the training is the same as the one being claimed. Quizzes can be offered at the end of the training to evaluate the knowledge of the trainees. Once staff members have successfully completed the training, a certificate of completion can be generated and stored in their account and a central database for future reference if needed.

Security reminders, protection from malicious software (malware), login monitoring, and password management are the four addressable implementation specifications. Each of these addressable implementation specifications has a single requirement (if reasonable

and appropriate for the CE), but the CE may choose from many options to meet the requirement. For instance, in the security reminder implementation specification, the requirement mandates the CE to implement periodic security updates. The organization can use print, electronic, or face-to-face reminders as the method to achieve this mandate. Security awareness and training specific to a workforce member's role in protecting the CE's information assets from potential damage resulting from malicious software is addressable. Monitoring system logins is used to help detect and prevent unauthorized access to information and includes such techniques as locking out users who enter their login name incorrectly more than once. In the password protection implementation specification, CEs must have appropriate policies and procedures in place to create, change, and safeguard passwords.

Security Incident Procedures Standard

The security incident procedures standard states that CEs must "implement policies and procedures to address security incidents" (45 CFR 164.308(a)(6)). The numerous preceding standards are all foundational supports that enable the achievement of this standard. Response and reporting is the single required implementation specification that states CEs must "identify and respond to suspected or known security incidents; mitigate, to the extent practicable, harmful effects of security incidents that are known to the covered entity; and document security incidents and their outcomes" (45 CFR 164.308(a)(6)). Examples of security incidents include such things as lost or stolen passwords, information system virus attacks, and theft of electronic media storage devices.

Contingency Plan Standard

The contingency plan standard is central to being able to ensure availability of data. This standard requires CEs to "establish (and implement as needed) policies and procedures for responding to an emergency or other occurrence (for example, fire, vandalism, system failure, and natural disaster) that damages systems that contain electronic protected health information" (45 CFR 164.308(a)(7)).

This standard is supported by three required and two addressable implementation specifications. The required data backup plan implementation specification requires CEs to "establish and implement procedures to create and maintain retrievable exact copies of electronic protected health information" (45 CFR 164.308 (a)(7)). To meet this implementation specification, organizations have to document what data needs to be backed up, and must understand the sources of all of those data. The disaster recovery plan implementation specification is also required and mandates CEs to "establish (and implement as needed) procedures to restore any loss of data" (45 CFR 164.308 (a)(7)). For the development of such a plan, a facility must determine the potential means of loss for their specific facility and location. Fire, water, and system destruction losses might be common among most organizations but loss due to tornadoes and hurricanes would be more specific to the geographic area. The required emergency mode operation plan implementation specification demands that CEs "establish (and implement as needed) procedures to enable continuation of critical business processes for protection of the security of electronic protected health information while operating in emergency mode" (45 CFR 164.308(a)(7)). The critical business processes must be identified as a first step before the procedures can be developed.

Testing and revisions procedures are the first of the two addressable implementation specifications in this section. Where this implementation specification is reasonable and appropriate for a CE, it must "implement procedures for periodic testing and revision of contingency plans" (45 CFR 164.308(a)(7)). This implementation specification applies to data backup plans, disaster recovery plans, and emergency mode operations. Many factors could influence the frequency and comprehensiveness of the testing and revisions

procedures. Some of these factors include the organization's size, as well as the availability of financial and human resources. Application and data criticality analysis is the second addressable implementation specification falling under this standard. When this implementation specification is deemed reasonable and appropriate for the specific CE, then it must "assess the relative criticality of specific applications and data in support of other contingency plan components" (45 CFR 164.308(a)(7)). In essence, this specification compels organizations to truly understand the value of their systems and to balance recovery and management efforts to the criticality level of each system in the organization. Procedures must also be in place for assessing the systems as new ones are added or any other changes are made in the normal operations of the systems.

Evaluation Standard

The eighth standard of the administrative safeguards is referred to as the "evaluation" standard (45 CFR 164.308(a)(8)). This standard is singular in requirement and has no supporting implementation specifications. It requires CEs to

> perform a periodic technical and nontechnical evaluation, based initially upon the standards implemented under this rule and subsequently, in response to environmental or operations changes affecting the security of electronic protected health information, that establishes the extent to which an entity's security policies and procedures meet the requirements of this subpart of the Security Rule. (45 CFR 164.308(a)(8))

The constant advances in technology and the growing use of new technologies are just two reasons demonstrating the need for this standard. An ongoing evaluation process is most likely to meet the spirit of this regulation, but the complexity of integrated health information technology often limits such an approach. The evaluation standard recognizes the limitations organizations may face in using an ongoing evaluation tool and supports alternative approaches provided they meet the intent of this standard.

Business Associate Contracts and Other Arrangements Standard

45 CFR 164.308(b)(1) is the final standard in the administrative safeguards section of the Security Rule; it addresses Business Associates, and it specifically discusses business associate contracts and other arrangements standards. By definition, a BA is, according to the HIPAA Privacy Rule, (1) an individual (or group) who is not a member of a CE's workforce but who helps the CE in the performance of various functions involving the use or disclosure of individually identifiable health information, or (2) a person or organization other than a member of a CE's workforce that performs functions or activities on behalf of or affecting a CE that involve the use or disclosure of individually identifiable health information (AHIMA 2017).

In January 2013, a final omnibus rule went into effect to provide greater protections for patient information than were in the original HIPAA of 1996 (OCR 2016a). The changes were, in part, to address items from the American Recovery and Reinvestment Act (ARRA) of 2009 and the Genetic Information Nondiscrimination Act (GINA) of 2008. One of the four final rules was to "make business associates of covered entities directly liable for compliance with certain of the HIPAA Privacy and Security Rules' requirements" (OCR 2016a). The final rule modified BAs to include health information organizations (HIO), e-prescribing gateways, and other persons who facilitate data transmissions, as well as vendors of personal health records.

Specific items that were addressed included that BAs must

- Follow the Security Rule for ePHI.

- Have business associate agreements (BAAs) with their subcontractors, who must also follow the security rule for ePHI. CEs do not have BAAs with these subcontractors.
- Obtain authorization prior to marketing (AHIMA 2013).

There is much implied under the first item regarding following the Security Rule. This means that the BAs must follow all of the requirements of the rule just as the CEs do. They will need to have their security procedures in place, assess them, have BAAs with their subcontractors, and report breaches or potential breaches of information, to name a few of their new responsibilities.

Check Your Understanding 11.1

Answer the following questions.

1. An inherent weakness or absence of a safeguard that could be exploited by a threat is a
   a. Security incident
   b. Breach
   c. Vulnerability
   d. Threat
2. The Security Rule safeguards are to address:
   a. All ePHI
   b. All patient records, paper and electronic
   c. The records being exchanged electronically
   d. Workforce working with patient records

3. Which one of the following is an administrative safeguard action?
   a. Facility access control
   b. Documentation retention guidelines
   c. Maintenance record
   d. Media reuse
4. The data backup plan requires that organizations
   a. Know what data needs to be backed up and sources of those data
   b. Create nightly data backup procedures
   c. Use off-site storage
   d. Act in accordance with the contingency plan
5. Which of the following is not a requirement of risk analysis?
   a. A full evaluation of the methods
   b. A full evaluation of the operational practices
   c. A full evaluation of the policies
   d. A full evaluation of the standard process

Indicate whether each of the following is true or false.

6. The Security Rule specifies the methods for conducting the required risk analysis.
7. The required implementation specification for the security incident procedures standard is response and reporting.
8. Addressable implementation specifications are optional for the facility to implement.
9. In a security risk analysis, the entity does not have to be concerned with the external sources of ePHI.

Physical Safeguards Standards

The Security Rule defines **physical safeguards** as "physical measures, policies, and procedures to protect a covered entity's electronic information systems and related buildings and equipment, from natural and environmental hazards, and unauthorized intrusion" (CMS 2007b). These standards are far reaching, as they require CEs to consider all physical access to the organization's ePHI. Like the administrative safeguards standards, the physical safeguards standards also have addressable and required implementation specifications. Where a CE deems an implementation specification to be reasonable and appropriate for their organization, they must then fulfill the addressable or required specifications.

Facility Access Control Standard

The facility access control standard requires CEs to "Implement policies and procedures to limit physical access to its electronic information systems and the facility or facilities in which they are housed, while ensuring that properly authorized access is allowed" (45 CFR 164.310(a)(1)). Important to this discussion is an understanding of what the term "facility" means. For the purposes of the Security Rule, the definition of facility is "the physical premises and the interior and exterior of a building" (45 CFR 164.310(a)(1)). The four addressable implementation specifications for this section of the rule are contingency operations, security plan, access control and validation procedures, and maintenance records.

The contingency operation addressable implementation specification requires CEs to "establish (and implement as needed) procedures that allow facility access in support of restoration of lost data under the disaster recovery plan and emergency mode operations plan in the event of an emergency" (45 CFR 164.310(a)(1)). Contingency operations are complex and will vary from organization to organization, with the intent being that CEs have the flexibility to determine the best approach for their given situation.

The facility security plan is the second addressable implementation specification under this standard. The intent of this implementation specification is that the facility security plan is to document the use of access controls. The CE is required to "implement policies and procedures to safeguard the facility and the equipment therein from unauthorized physical access, tampering, and theft" (45 CFR 164.310(a)(1)). To meet this implementation specification, organizations should consider methods such as lock-and-key controls, security tagging equipment, using video cameras for surveillance, monitoring identification badges, and the use of human workforce to perform facility security checks.

The third addressable implementation specification associated with this standard is access controls and validation procedures. This specification indicates that CEs must "implement procedures to control and validate a person's access to facilities based on their role or function, including visitor control, and control of access to software programs for testing or revision" (45 CFR 164.310(a)(1)). The characteristics of the CE (for example, size) could influence some of the decisions made regarding this specification. For example, in large facilities, badge access might be visually confirmed for every workforce member every day. In smaller organizations where the staff is small, identification badge access confirmation might not be a daily activity because, presumably, everyone knows the workforce members by face. In addition to workforce members, CEs must evaluate practices to control and limit visitor access to information. For instance, the computer screens used by healthcare providers (such as nurses and doctors) in a hospital are placed so that they face off the hallway. This implementation specification requires CEs to document the rationale for how they reach their security decisions.

Maintenance records are the final implementation specification in this section. This specification states that CEs must "implement policies and procedures to document repairs and modifications to the physical components of a facility which are related to security (for example, hardware, walls, doors and locks)" (45 CFR 164.310(a)(1)). Again, environmental (facility size, location, and so on) factors can influence how an organization makes its decisions for this specification. Manual logs might be appropriate in some cases and more sophisticated controls might be appropriate in other cases.

Workstation Use Standard

The workstation use standard functions without additional implementation specifications. To meet this standard, the CE must

> Implement policies and procedures that specify the proper functions to be performed, the manner in which those functions are to be performed, and the physical attributes of the surroundings of a specific workstation or class of workstation that can access electronic protected health information. (45 CFR 164.310(b))

An important note is that this standard applies not only to workstations located physically in the facility. It applies at a much broader level to workforce members using workstations off-site to complete their work activities. At a minimum, any safeguards that are required in the office must also be required off-site. An example of this kind of safeguard are timeouts or logouts that break the system connection if the workstation remains idle for a specified period of time, for instance, 10 minutes.

Workstation Security Standard

The workstation security standard requires CEs to "implement physical safeguards for all workstations that access electronic protected health information, to restrict access to authorized users" (45 CFR 164.310(c)). This standard differs from the previous workstation use standard that addresses policies and procedures for how workstations should be used and protected. The workstation security standard addresses how workstations are to be physically protected from unauthorized access.

Device and Media Controls Standard

The device and media controls standard states that CEs are to "implement policies and procedures that govern the receipt and removal of hardware and electronic media that contain electronic protected health information, into and out of a facility, and the movement of these items within the facility" (45 CFR 164.310(d)(1)). Here electronic media are defined as "electronic storage material on which data is or may be recorded electronically, including, for example, devices in computers (hard drives) and any removable/transportable digital memory medium, such as magnetic tape or disk, optical disk, or digital memory card or transmission media used to exchange information already in electronic storage media" (45 CFR 160.103).

A major consideration for organizations trying to meet this standard is the need to understand what all their information resources are and where they all lie. This often leads to the need for establishing an effective systems inventory method. Not only do organizations need to document their existing and new systems, they must also track movement of components of each system as well as track system and system component obsolescence and replacement. Two addressable and two required implementation specifications support this standard. The two addressable specifications are accountability and data backup and storage. The required specifications are disposal and media reuse.

The required disposal implementation specification states that CEs must "implement policies and procedures to address the final disposition of electronic protected health

information, and/or the hardware or electronic media on which it is stored" (45 CFR 164.310(d)(2)). There are multiple methods that can be used to make ePHI unusable or unreadable. The Security Rule specifically mentions degaussing as one of those methods, but stops short of requiring degaussing as the preferred method for rendering data unreadable or unusable when a CE wishes to dispose of electronic media that contains ePHI.

The required media reuse implementation specification states that CEs must "implement procedures for removal of electronic protected health information from electronic media before the media are made available for re-use" (45 CFR 164.310(d)(2)). This implementation specification is applicable to both internal and external reuse of media scenarios and requires organizations to develop policies and procedures to address media reuse.

The addressable accountability implementation specification states that the CE must "maintain a record of the movements of hardware and electronic media and any person responsible therefore" (45 CFR 164.310(d)(2)). This implementation specification presents a growing challenge in many healthcare organizations where mobile technology is increasing in use, increasing in cost, increasing in data storage, and decreasing in size. The decreasing size factor creates an increased risk due to the opportunity for concealment and theft. The flipside to this discussion is the value it offers in terms of user convenience. Like so many concepts addressed by HIPAA, organizations must devise methods to ensure an appropriate balance of protection and access or use. This addressable implementation specification speaks again to the need for inventorying and tracking of electronic media.

Data backup and storage is the final addressable implementation specification for this standard. Under this implementation specification, the CE must "create a retrievable, exact copy of electronic protected health information, when needed, before movement of equipment" (45 CFR 164.310(d)(2)). This implementation specification parallels the data backup plan for the contingency plan standard in the administrative safeguards section of the rule. As such, both components of the rule are often fulfilled with a combined policy and procedure that meets the requirements of both sections.

Technical Safeguards Standards

As technological advances occur, they often bring with them new security challenges. The conundrum of duality of interests again presents itself here. Advancing technology leads to great opportunity, but at the same time increases organizational risk. Balancing these two facets to capitalize on the benefits and reduce the risks is often at the center of both the HIPAA Privacy and Security Rules. Therefore, there are technical safeguards in the Security Rule.

The Security Rule defines **technical safeguards** as "the technology and the policy and procedures for its use that protect electronic protected health information and control access to it" (45 CFR 164.304). In keeping with the flexibility, scalability, and technological neutrality concepts of meeting the Security Rule requirements, CEs must determine which security measures and technologies are reasonable and appropriate for implementation in its organization (CMS 2007c).

Access Control Standard

The access control standard directs CEs to "implement technical policies and procedures for electronic information systems that maintain electronic protected health information to allow access only to those persons or software programs that have been granted access rights as specified in §164.308(a)(4) [information access management]" (45 CFR 164.312

(a)(1)). Of importance here is the concept that access controls should be appropriate for the role and function of the workforce member. Two required and two addressable implementation specifications are included in this standard.

The required unique user identification implementation specification states that CEs must "assign a unique name and/or number for identifying and tracking user identity" (45 CFR 164.312(a)(1)). Unique identification methods allow for the user's actions to be monitored and serve as the method of enforcing accountability of user actions.

The second required implementation specification is referred to as the emergency access procedure. This specification states that CEs must "establish (and implement as needed) procedures for obtaining necessary electronic protected health information during an emergency" (45 CFR 164.312(a)(1)). This description, at face value, conveys the importance of ensuring availability of information as needed to perform one's role in emergency situations. From a technology standpoint, many systems have a "**break the glass**" method that allows users to access data that they may not otherwise be allowed to access. The users are alerted in advance that their actions are being monitored, giving them an opportunity to halt their actions, if inappropriate.

Automatic logout is the first addressable implementation specification, which states that CEs must "implement electronic procedures that terminate an electronic session after a predetermined time of inactivity" (45 CFR 164.312(a)(1)). This concept is probably not foreign to readers, as it is commonly used in many types of information systems, healthcare and otherwise.

Encryption and decryption constitute the final addressable implementation specification. Where this implementation specification is a reasonable and appropriate safeguard for a CE, the CE must "implement a mechanism to encrypt and decrypt electronic protected health information" (45 CFR 164.312(a)(1)). **Encryption** is a technical method that reduces access and viewing of ePHI by unauthorized users. It is defined as "the use of an algorithmic process to transform data into a form in which there is a low probability of assigning meaning without use of a confidential process or key" (45 CFR 164.304). To perform encryption, one needs to have three items: the data to be encrypted, the chosen encryption algorithm, and an encryption key corresponding to the encryption algorithm. After the data are encrypted, the meaning of the message is not obvious without decryption, which is the reverse process of encryption. There are different encryption algorithms using different keys. Their protection strength, or the difficulty (in terms of time and efforts) in figuring out the meaning of the message without knowing the key for decryption, are different. The Security Rule does not have specific requirements about encryption algorithms, but in practice, strong encryption algorithms (for example, Advanced Encryption Standard, or AES, with a 256-bit key) are typically used in health information systems to protect ePHI.

Audit Control Standard

The audit control standard is singular in focus and requires CEs to "implement hardware, software, and/or procedural mechanisms that record and examine activity in information systems that contain or use electronic protected health information" (45 CFR 164.312(b)). This standard recognizes that, regardless of the controls put into place to prevent unauthorized access, there must be a way to track and record user activities in the system. Such tracking can be used to monitor intentional and unintentional actions taken by users to access ePHI.

Integrity Standard

The integrity standard is supported by one addressable implementation specification. The integrity standard states that CEs must "implement policies and procedures to protect

electronic protected health information from improper alteration or destruction" (45 CFR 164.312(c)(1)). This standard is intended to ensure data integrity, which is defined as the extent to which healthcare data are complete, accurate, consistent, and timely; it is a security principle that keeps information from being modified or otherwise corrupted either maliciously or accidentally. Ultimately, data integrity supports high-quality healthcare, and the effect of compromised data integrity can be significant. For these reasons, organizations must take steps to ensure data are not improperly altered or destroyed. The addressable implementation specification mechanisms to authenticate ePHI requires CEs to "implement electronic mechanisms to corroborate that electronic protected health information has not been altered or destroyed in an unauthorized manner" (45 CFR 164.312(c)(1)).

Person or Entity Authentication Standard

The person or entity authentication standard requires CEs to "implement procedures to verify that a person or entity seeking access to electronic protected health information is the one claimed" (45 CFR 164.312(d)). Simply put, this standard seeks to ensure that organizations put methods in place to verify that users are who they claim they are. Passwords, passphrases, smart cards, tokens, and biometrics (for example, fingerprints) are some of the many methods used in healthcare settings to confirm user identity. To strengthen authentication, in recent years, multifactor authentication has been widely adopted in healthcare. For instance, after the username and password are entered, the user is required to enter a one-time passcode generated on a mobile device, such as a smartphone.

Transmission Security Standard

The transmission security standard recognizes the potential risk involved with data in transit and requires CEs to "implement technical security measures to guard against unauthorized access to electronic protected health information that is being transmitted over an electronic communications network" (45 CFR 164.312(e)(1)). Two addressable implementation specifications are associated with this standard.

The integrity controls implementation specification directs CEs to "implement security measures to ensure that electronically transmitted electronic protected health information is not improperly modified without detection until disposed of" (45 CFR 164.312(e)(1)). Specifically covered in this implementation specification is the assurance that ePHI is not improperly modified during transmission. To detect whether ePHI has been modified during the transmission, one approach is to use a hash function, which would generate different hash values if two data files have one single-bit difference.

The encryption addressable implementation specification states that CEs must "implement a mechanism to encrypt electronic protected health information whenever deemed appropriate" (45 CFR 164.312(e)(1)). Encryption converts the original message into an encoded or unreadable message. When decrypted, the message returns to its original state and can be read by the recipient. If a sender uses one algorithm (for example AES) to perform encryption, the receiver must use the corresponding decryption algorithm (AES) and the correct key to perform the desired decryption. In some encryption/decryption algorithms, the same key is used by both processes. In some algorithms, the encryption and decryption keys are different.

To meet the requirements for transmission security, transmitted ePHI that has been inappropriately altered is detected and destroyed, and ePHI is encrypted where necessary. The following are a few examples of how ePHI can be inappropriately made available to unauthorized individuals:

- An unencrypted portal hard drive containing patient information was stolen from a hospital.
- A hospital employee created an unencrypted, not password-protected patient list on the employee's personal laptop to help with her job of cataloguing records.
- Information was compromised for 8,800 members of a health plan when an email including PHI was inadvertently sent to an incorrect email address. The PHI was not encrypted.

The potential for breaches with transmission of PHI are wide and varied. In addition to the technical requirements to protect against this kind of breach, policies, procedures, and employee education are very important.

Confidentiality, Integrity, and Availability

The HIPAA Security Rule seeks to uphold confidentiality, integrity, and availability of data. The previously discussed administrative, physical, and technical safeguards also require and recommend methods to ensure confidentiality, integrity, and availability of data.

Confidentiality

Confidentiality is "a legal and ethical concept that establishes the healthcare provider's responsibility for protecting health records and other personal and private information from unauthorized use or disclosure" (AHIMA 2017, 34). In the context of the Security Rule, confidentiality means that ePHI is accessible only by authorized people and processes. Encryption is the typical way to protect the confidentiality of ePHI.

Integrity

Maintaining data integrity is central to the Security Rule. Under the Rule, **integrity** means that ePHI is not altered or destroyed in an unauthorized manner. With continuing advancements in technology and electronic data management come additional legal obligations for ensuring data integrity. In addition to Security Rule requirements, concepts such as e-discovery that implore organizations to establish and follow methods to ensure integrity are becoming commonplace in healthcare organizations.

Availability

Equally as important as protecting ePHI from unauthorized users is the concept of ensuring that it can be accessed as needed by authorized users. **Availability** refers to "the property that data or information is accessible and useable upon demand by an authorized person" (CFR 45 164.304). That is, patient information must be available for treatment and other vital uses. Encompassed in the concept of availability are many of the previously discussed safeguards, such as the data backup and storage specification. The wide adoption of cloud-based services can improve ePHI availability as well.

Penalties and Enforcement

Penalties and enforcement were addressed in the original HIPAA bill but have changed over time. The determination of penalties is complex depending on the kind of violation; whether it was an inadvertent release of information, neglectful, or done knowingly; whether it was a primary or subsequent violation; and whether there was negligence in correcting known policies and procedures that led to violations. The history of the enforcement is complex due to the numerous changes of the agencies responsible;

however, the duplication of efforts has been recognized and responsibility streamlined as a result.

The Office for Civil Rights

The **Office for Civil Rights (OCR)** is currently the enforcement agency for the HIPAA Privacy and Security Rule. Since April 14, 2003, the Privacy Rule has been enforced by the OCR. The Security Rule was enforced by the Centers for Medicare and Medicaid Services (CMS) from its effective date of April 20, 2005, through July 26, 2009. Since that time, the Security Rule has been enforced by the OCR as well. The OCR must investigate all reported violations and appropriately initiate investigations for cause in absence of a reported violation.

page 317

Civil Penalties

The Privacy and Security Rules both outline various penalties for violations. **Civil penalties** are generally fines or money damages used to sanction violators. The Final Rule for HIPAA modified civil penalty enforcement language in both the Privacy and Security Rules by outlining a tiered structure of enforcement guidelines. The tiered method specifies a minimum to maximum penalty range based on the specific category of violation.

The four main categories of civil violations escalate in nature and start at the bottom rung with violations where the individual did not know (or through reasonable diligence would not have known) that he or she violated HIPAA, and range to the fourth and final category, which is reserved for those who violate HIPAA through willful neglect and do not correct the violation pattern. The civil monetary penalties prior to February 18, 2009, were restricted penalties of not more than $100 per violation, or not in excess of $25,000 for identical violations during a calendar year. The amount of civil monetary penalties has been adjusted multiple times since February 18, 2009. From 2022, HIPAA violations are subject to penalties of up to $60,226 per violation and up to $1,806,757 per calendar year for multiple violations of the same HIPAA provision. If one organization has violations of multiple HIPAA provisions occurring over multiple years, the penalties can be much higher than $1.8 million (HHS 2021).

Criminal Penalties

If the OCR reviews a complaint and determines that it qualifies as a criminal violation, the case is referred to the Department of Justice (DOJ). In 2005, the DOJ clarified that criminal violation enforcement is applicable to both individuals and CEs. When **criminal penalties** are appropriate, the OCR works in conjunction with the DOJ to pursue possible violators. Criminal penalties can include imprisonment of 1 to 10 years, a fine, or both, plus possible restitution of monies received. When the breach of PHI is found to be a HIPAA violation, the criminal violations vary by cause beginning with (1) reasonable cause or possibly being unaware of the violation, and continuing with (2) acquiring PHI by false pretenses, and (3) obtaining the PHI for personal gain or malicious intent, which has the most severe penalties (45 CFR 160.404).

Enforcement Activities to Date

The **HIPAA Enforcement Rule** is found at 45 CFR 160, subparts C, D, and E. This rule spells out the authority of the OCR as previously described. The complaint process is fully outlined on the OCR website and presents the process from initial complaint through referral, investigation, and resolution. After investigation, resolutions range from the OCR finding no violation to obtaining corrective action to the issuance of formal findings of violation. In addition, reported violations could be found unenforceable due to

page 318

technical issues such as not filing the complaint within the required time frame, the entity not being considered a CE, or the incident as described not violating the rules.

In 2020, there were 26,530 complaints filed with the OCR. Among those complaints, a large portion of them (19,826; 74.7 percent) were resolved after intake and review. OCR closed these cases without further investigation. This might happen, for example, when the breached data do not have ePHI included, or the entity that violated the HIPAA rules is not a CE or BA. For 5,091 complaints, OCR provided technical assistance to the CE, BA, and complainant through early intervention by OCR investigators. There was no further investigation on these complaints, either. The remaining 1,613 complaints were investigated. Of those, in 661 cases it was determined that no violation occurred. OCR provided post-investigational technical assistance to 80 entities, and required 872 entities to make changes according to HIPAA regulations. The situation in recent years was similar to that in 2020 (HHS 2020).

The following case study is an example of an issue with the HIPAA Privacy and Security Rules. The outcome of this case was a resolution agreement, a contract between HHS and the covered entity outlining obligations to be met, such as the development of policies and procedures or required staff training, plus the required reporting to HHS for a specified period of time.

Case Study: HIPAA Privacy and Security Rule Violation

In violation of the HIPAA Privacy and Security Rules, two entities within a healthcare system, home health services and hospice services, had several instances over about a seven-month period where backup tapes, optical disks, and laptops were left unattended and eventually lost or stolen. All of the devices contained unencrypted ePHI on more than 386,000 patients. The healthcare organization responded to state notification laws and to the HHS with more than 30 complaints submitted as a result. Working together, OCR and CMS focused the investigation on the organization's failure to implement policies and procedures to safeguard this information. The HHS entered into a resolution agreement with the organization to settle the violation. It was agreed that the CE would pay a $100,000 resolution amount (due to cooperation, this resolution amount was in lieu of a civil monetary penalty that could have been much more) and implement a robust corrective action plan. This plan included revising policies and procedures for physical and technical safeguards (encryption), governing off-site transport and storage of the electronic media that contained PHI, training workforce on the safeguards, conducting routine audits and site visits of facilities within the organization, and submitting the required compliance reports to HHS for three years.

This case is an example of a situation where a resolution agreement is suitable. In the majority of compliance issues investigated, there is one facility with one type of breach involved and one major corrective action imposed. In this case there were two different entities with several different types of data storage devices involved. There were several different corrective actions including a payment to be made and a comprehensive corrective action plan with numerous items to be accomplished over a period of time. The required compliance reports assured compliance and

accountability.

Breach Notification

In January 2013, the Final Rule on breach notifications went into effect. The Final Rule is based on the original HIPAA rules and supplants an interim final rule published in 2009. It fulfills provisions of the HITECH Act and strengthens privacy and security protection of PHI. The Final Rule increased patients' privacy protections, provides patients with new rights regarding their health information, and increases the ability for enforcement of the law to protect PHI (OCR 2013).

The rule requires CEs and BAs to report breaches of "unsecured protected health information. Unsecured protected health information is protected health information that has not been rendered unusable, unreadable, or indecipherable to unauthorized individuals through the use of a technology or methodology" (ARRA 2009). Protecting health information from unauthorized use is accomplished by using methods that will make it unusable, unreadable, or indecipherable (HHS 2016). NIST has issued a series of special publications that address encryption of data at rest and in motion as well as destruction of both electronic and nonelectronic media. These special publications are numbered in the 800 series and can be located on the NIST website (NIST 2016).

Reporting requirements for breach mandate notification to the individual whose information was breached, and in the case of breaches of more than 500 individuals' information, to the media and the Secretary of HHS. There are three exceptions:

- The first exception is applicable when a workforce member unintentionally acquires, uses, or discloses PHI while acting under the authority of a CE or BA.
- The second exception is applicable when an authorized workforce member inadvertently discloses PHI to another authorized workforce member within the same CE or BA setting.
- The final exception is applicable when the CE or BA who made an inadvertent disclosure has reason to believe that the recipient of the PHI would not have been able to retain the information. For example, if the provider called a patient with a common name to discuss laboratory results, they might dial the wrong phone number. Although they might discuss "John Smith's" lab results, as long as no hard-copy results were sent to the wrong patient, it is reasonable to believe that the wrong John Smith could not keep the information.

Risk Assessment

CEs and BAs must have procedures in place to conduct a risk analysis. This is done to assess the potential risks and areas of vulnerability to the security of the PHI just as is required for the privacy aspect of HIPAA. It is important to identify and correct those things that might threaten the confidentiality, integrity, and availability of the information. Failure to conduct such an audit could be detrimental in the event of an audit or investigation of a complaint.

Check Your Understanding 11.2

Answer the following questions.

1. A systems inventory does not include
   a. Existing and new systems
   b. Movement of components of each system

c. Tracking system and system component obsolescence
d. Training records for employees to be able to use the systems

2. The technology and the policies and procedures for its use that protect electronic protected health information and control access to it are
   a. Administrative safeguards
   b. Technical safeguards
   c. Physical safeguards
   d. Integrity controls
3. The integrity standard protects
   a. From unauthorized access
   b. From unauthorized use
   c. From improper transmission of data
   d. From improper alteration or destruction
4. Which of the following is not a resolution to HIPAA violations?
   a. No violation found
   b. Obtaining corrective action
   c. Issuance of formal findings of violation
   d. No investigation required

Indicate whether each of the following is true or false.

5. When encrypted information is decrypted, it is in a different format than the original state.
6. Maintaining access to ePHI is less important than the protection of ePHI.
7. The difference between the levels of civil penalties has to do with whether the CE knew or should have known that a violation occurred.

Medical Identity Theft and Red Flag Rules

Health information serves many functions, the most important of which is patient care. However, health information serves many other functions. The legal considerations for health information extend beyond privacy, security, and the legal EHR. The legal considerations now include the responsibility of providers to protect against medical identity theft, consider legal liability related to medical malpractice in light of EHR use, and incorporate disaster planning and preparedness. The first two issues are relatively new and the third requires new and different ways of thinking, but they are vitally important in the evolving world of health information technology and EHRs. This section discusses the challenges and recommendations for addressing each of these.

Medical Identity Theft

Medical identity theft occurs when personal information (such as name, Social Security number, or Medicare number) is taken without the patient's knowledge to obtain medical care, buy drugs, or submit fake billings to Medicare in the name of the patient. Medical identity theft can disrupt a person's life, damage credit rating, and waste and inappropriately use taxpayer dollars. The damage can also alter treatment or even be life threatening to a victim of identity theft if the wrong information subsequently ends up in the patient's health records (HHS, Health IT 2015).

There are implications for the patient when health records are accessed and modified, credit accounts are created, and credit scores are accessed and changed. There are also consequences for the third-party payers related to their payment for services, drugs, and equipment. There are costs that the facility or provider must bear associated with follow-

up to resolve claims and receive the appropriate payment. The patient health information must also be rectified so there will be complete and correct information on each patient for continuity of care and to alleviate issues in releasing PHI when information on the wrong person is combined with that of the patient.

Instances that would not be considered identify theft are where health information on the wrong patient was inadvertently put in the incorrect record. When the financial information on a patient is used to purchase nonmedical items and services, this is not considered medical identity theft. The key factors for a situation to be medical identity theft are that a form of a patient's identification is taken without the patient's knowledge and the identification is then used to obtain some sort of medical services.

In addition to the financial implications of medical identity theft, the theft carries the very real possibility that the incorrect information could be used in the delivery of patient care, putting patients at risk. This can occur if critical medical conditions, procedures, medications, allergies, and other information are either omitted from the record or wrongfully included. This would occur, for example, when a medical report of the perpetrator is filed in the health record of a patient and the report indicates the blood type to be different than that of the victim of the identity theft. This might not be noticed for some time, if ever, because the erroneous report would have the patient's name and other identifying information filed among reports with the same identifying information. Unfortunately, victims of identity theft could find themselves unable to access their records because some providers will refuse to release the records containing another patient's information—even if the other patient broke the law.

Red Flag Rules

In response to increasing identity theft, often used for financial gain, the Federal Trade Commission (FTC), Department of the Treasury, Federal Reserve System, Federal Deposit Insurance Corporation, and the National Credit Union Administration simultaneously issued regulations known as the Red Flag Rules in November 2007 (King and Williams 2008). These rules require creditor and financial institutions to implement an Identity Theft Prevention Program (King and Williams 2008). The portion of these rules that apply to healthcare organizations are contained in 16 CFR 681, under the enforcement authority of the FTC (Springer 2009). The **red flags** that are addressed in the rules are the potential patterns, practices, or specific activities indicating the possibility of identity theft (FTC 2013). Examples of red flags include inconsistencies in documents (driver's license and insurance card), documents that appear to be altered, and not having more than one form of ID (insurance card but no other form of ID).

First, healthcare organizations must understand whether they are subject to the Red Flag Rules. Healthcare providers may be subject to the Red Flag Rules if they (1) meet the definition of creditor under the Fair Credit Reporting Act; or (2) use consumer credit reports, subjecting them to the Address Discrepancy Rule (Springer 2009). A healthcare creditor would include any organization that allows for payment on medical services provided to a patient after those services were provided or over a period of installment payments (King and Williams 2008). This includes most healthcare organizations, especially hospitals. Fortunately, the Red Flag Rules include guidelines to assist in the development of an identity theft prevention program:

1. Identify Covered Accounts: Most patient accounts and billing records qualify as covered accounts because they contain information sufficient to enable identity theft if lost or stolen.
2. Identify Relevant Red Flags: A red flag is a pattern or practice of specific activity that indicates the possible existence of identity theft. Red flags can include alerts,

notifications, or other warnings that might be received from consumer report agencies or other service providers; suspicious documents such as identification cards or other documents that are inconsistent or appear to be falsified; other suspicious information such as odd changes of address; or other unusual use of or other suspicious activity related to a given account.

3. Detect Red Flags: Organizations must develop reasonable approaches for detecting potential red flags. This requires that organizations determine what information and documentation will be required of patients. Many now require both an insurance card and a photo identification card such as a driver's license. Additionally, organizations must establish policies and procedures. This is quite important because liability increases if the workforce either cannot or does not follow the policies and procedures. Thus, the policies and procedures should support the detection of red flags in a feasible way.
4. Respond to Red Flags: An adequate response to red flags includes prevention and mitigation via monitoring covered accounts, contacting patients or consumers if needed, changing passwords and security codes, handling accounts carefully, or notifying law enforcement. These responses will, of necessity, be context-dependent and often require judgment on the part of the healthcare organizations. For example, what will the organization do if the patient cannot produce all the identification required or the identification presented has discrepancies?
5. Oversee the Program: As with all operations, the identity theft prevention program falls under the purview of the board of directors, trustees, or the designated member of management.
6. Train Employees: The workforce, especially those responsible for creating, maintaining, and administering patient accounts, must be trained in all aspects of identity theft prevention.
7. Oversee Service Provider Arrangements: The healthcare organization must take steps to ensure that any service providers granted access to covered accounts carry out the identity theft prevention program.
8. Approve the Identity Theft Prevention Program: The board of directors or appropriate committee must review and approve the identity theft prevention program.
9. Provide Reports and Periodic Updates to the Identity Theft Prevention Program: A written report on the status of the program, service provider arrangements, any significant incidents, and recommendations for any needed changes should be included in the annual report. (King and Williams 2008)

Organizations that have already implemented the Red Flag Rules report finding increases in attempted and successful identity theft (Keith 2010). Unfortunately, they report that it is difficult to detect red flags at the time identity theft is being perpetrated, although they do try. Some have implemented special designations in their registration pathways so that all employees can identify records of victims, perpetrators, or those with suspicious activity. Yet, most medical identity theft is still detected when the bill is issued (Keith 2010).

Operational Recommendations

The obligations of health informatics (HI) and HIM professionals related to medical identity theft fall into three general areas. The first is to urge and educate consumers to adopt preventive measures such as these:

- Exercising caution when sharing personal and health information with providers
- Monitoring the Explanation of Benefits (EOB) they receive from insurance companies
- Maintaining copies of healthcare records
- Monitoring credit reports and history documents for unexpected medical charges or liens
- Protecting all health insurance and financial information (AHIMA e-HIM Work Group on Medical Identity Theft 2008).

The second area is establishing organizational methods to prevent and detect medical identity theft. Some of these methods include the following.

- Conducting and acting on an information security risk analysis on an annual basis
- Ensuring background checks are performed for employees and BAs when hiring, especially for work in high-risk areas
- Establishing patient identification verification processes that are compliant with the HIPAA Security Rule, but are also feasible for high-throughput areas with significant time demands
- Minimizing the use of Social Security numbers for identification, suppressing them when possible and, when not, displaying only the last four or six digits of the numbers
- Implementing policies and procedures that safeguard the privacy and security of individually identifiable health information, in compliance with applicable HIPAA and other federal and state regulations
- Creating an alerting and identity theft response plan if any suspicious activity is discovered
- Implementing ongoing staff training programs to ensure all employees with access to identifiable patient information can assist in preventing identity theft (AHIMA e-HIM Work Group on Medical Identity Theft 2008).

The third medical identity theft area where HIM professionals have an obligation to assist is related to the data in the patient record; this is particularly important. The steps an HIM professional can help with include the following:

- Drafting policies and procedures to ensure victims of medical identity theft are not denied access to their patient records
- Establishing mechanisms to correct inaccurate information in patient records, which includes assisting victims in identifying those who might possess inaccurate records by providing a full accounting of disclosures
- Staying abreast of medical identity theft–related legislation and regulations
- Providing victims of medical identity theft with a checklist of actions and resources. These are some examples of action items included in the checklist:
  - Review credit reports, correct them, and place a Fraud Alert on them.
  - For stolen passports, contact the US Department of State.
  - Request an accounting of disclosures of medical records.
  - Review health records to assure that information has been corrected prior to seeking healthcare (AHIMA e-HIM Work Group on Medical Identity Theft 2008).

Medical identity theft and financial identity theft cause significant problems for consumers, insurers, and providers at a minimum. It is important that all possible steps be taken to both prevent medical identity theft and mitigate or minimize the effects of medical identity theft once it occurs.

Disaster Preparedness

Hurricanes, tornadoes, flooding, earthquakes, and man-made disasters such as terrorist attacks and mass shootings cannot, unfortunately, be ignored by HI and information professionals. HI and HIM professionals have two very important roles related to disaster preparedness: (1) ensure the protection of the organization's information assets, and (2) ensure the information functions will continue in the event of a natural or man-made disaster affecting the entire organization. These two roles can be separated because the first can occur without the second, but they are related. The first role of protecting the information assets needs to be done in all situations, including those that might not technically be termed a disaster, such as attacks by hackers or technical system failure due to network or software malfunctions. Because this role focuses largely on protecting the information assets, it is largely under the purview of the HIM and IT departments. The second role crosses all departments and functions in the organization and, although HIM and IT are expected to meet the organizational needs, it is extremely collaborative in nature.

Protecting Information Assets

Protection of organizational information assets is a part of the HIPAA Security Rule. Although the technical requirements of the Security Rule seem fairly straightforward, the information security team must prepare thoroughly, thinking outside of the box, envisioning worst-case scenarios or trying to imagine how a technically savvy person could do harm to the organization via the different installed information systems. NIST has published guides to help organizations, including the federal government, prepare contingency plans for information systems operations. NIST Special Publication 800-34 Rev. 1, "Contingency Planning Guide for Federal Information Systems," gives guidance and provides multiple templates for developing disaster plans for information services. NIST Special Publication 800-30, Rev. 1, "Guide for Conducting Risk Assessments," can also help with information asset protection planning. These and additional NIST guidance can be found at http://csrc.nist.gov/publications. Several examples of risks follow to illustrate the possibilities.

page 326

Significant risks to information systems that might put the organization at risk include hackers attacking computers that control medical equipment and facilities that are on networks with weak security, or significant downtime due to a massive system failure. For example, the heating, ventilation, and air conditioning system for a medical center had been set up to be accessed remotely by an outside company that managed the system (Gupta 2011). Hackers attempted to use this poorly secured computer to gain access to other systems. After investigation by law enforcement, it was discovered that the system had been attacked previously more than 10 times (Gupta 2011). In that same vein, another healthcare system experienced a worst-case scenario when the geographic area where their data system was located suffered a massive loss of power (Conn 2012). Because it was not where the organization itself was located, operations in the organization needed to continue unimpeded. Unfortunately, the backup generators only worked for a short time (Conn 2012).

IT and HIM personnel must prepare for these and other potential threats to their systems with the initial step of a **business impact analysis** (BIA). Here a BIA is one

method for evaluating and prioritizing the risks to help make decisions to best address the risks. Concepts important to a BIA are the **recovery point objective (RPO)** and the **recovery time objective (RTO;** Ranajee 2012). The RPO represents the length of time that you can operate without a particular application. The RTO is the maximum amount of time tolerable for data loss and capture (Ranajee 2012). The goal of the BIA is to identify any gaps in the current recovery capability and to develop a strategy for meeting the identified RTO and RPO. The following steps should be taken and the following questions should be asked:

1. Identify the minimal resources required to maintain business operations. What are essential operations and what is required to keep them running? What might be out of sight downstream that needs to be considered?
2. Determine the business recovery objectives and assumptions. How much downtime and loss of data can each department sustain? How are the data received and processed by each department?
3. Establish the order of priority for restoration of business functions. What are the key patient care departments? What are the IT applications that support these critical operations?
4. Estimate the operational, financial, and reputational impact due to loss of data. Although IT and HIM can provide data and information regarding these impacts, it is more appropriate for executives with input from the board of directors or trustees to make the final estimations and decisions (Ranajee 2012). Table 11.1 illustrates a matrix for conducting a BIA. This tool is used to identify the risk level and prioritize the potential threats identified.

Table 11.1 . Information security threat analysis quadrant

	Low risk to operations	High risk to operations
Low probability of occurring	Green	Orange
High probability of occurring	Yellow	Red

The colors used are similar to what is used in a heat map to show temperatures geographically, with red representing the hottest, then orange, yellow, and green representing the coolest. Items that fall into the red or orange quadrants should be handled first. Those that fall into the yellow and green quadrants should still be addressed but are a lower priority.

Once the BIA and prioritization have been accomplished, many decisions remain to be made, such as the type and location of backup data facilities. There are usually three different types of facilities:

1. **A hot site** is for critical applications and operations, which can be online within hours. This option is the most expensive.
2. **A warm site** provides basic infrastructure, but takes time, possibly a week, to activate. Obviously, this is less expensive.
3. **At a cold site**, equipment must be brought in, but the site is powered and secure. This type of facility is much less expensive, but could take up to a month to operationalize (Ranajee 2012).

Requirements such as HIPAA security regulations, BAAs, and so forth continue to pertain. Additionally, the backup data center should be located in a low-risk geographic location (that is, not in an earthquake or hurricane zone), and should have multiple layers

of physical security such as biometrics, mantraps, video monitoring, and so on; expansion capabilities; and an established history of uptime and redundancy (Ranajee 2012).

Protecting the various patient care and other critical systems, some of which might not have been considered (such as HVAC systems), is the first step to health information disaster preparedness. These systems can be attacked or compromised in situations that do not necessarily include natural or man-made larger disasters and must be planned for accordingly.

In recent years, cloud-based information services and mobile devices have been widely used in healthcare facilities. These facilities do not need to maintain their own servers and the associated network systems. The information is always available as long as there is power and internet access, as the cloud-based services, by their nature, are distributed in multiple geographical locations, which are unlikely to be affected by the same disaster at the same time. On the other hand, cloud-based services also have their own drawbacks. For instance, because the data are distributed in multiple places, the availability of the data is very good, but these multiple places involve different countries, making the enforcement of HIPAA Privacy and Security Rules highly problematic.

Continuity of Information Services

Keeping operations going during an external disaster, whether an act of nature or man-made crisis, is required. Although disasters often result in some type of gap in services, a significant failure to plan could be considered negligence in certain circumstances. This section offers recommendations for organization-wide information services **disaster recovery planning (DRP)**. For the purposes of this book, DRP for organizational information services is focused on the recovery from physical disasters, including some elements usually included as a part of **business continuity planning (BCP)**. BCP includes the recovery and use of the technology as in DRP, as well as the ability of the organization to continue the processes required for ongoing business operations. Sometimes BCP requires operational procedures in the absence of IT.

Overview of Disaster Planning

Disaster planning can occur at many different levels, especially in a healthcare organization. It can occur for the entire organization, but departments can also plan for disaster, as can the entire community, which could be dependent on that healthcare organization. In conjunction with many different levels of planning, the components of preparedness, planning, response, and recovery are involved in disaster planning.

One of the strategies learned after major disasters in the United States is that the workplace should help its workforce with their individual disaster planning. Given that the disaster plan for any organization often rests with the workforce, the organization might want to ensure that its "essential" personnel have a personal disaster plan in place. For example, in an earthquake zone, do the different family members know how they are supposed to communicate if an earthquake occurs? In a hurricane or flood zone have employees thought through when, where, and how their families would evacuate? If power is lost, does the employee have a generator or a backup plan? One study reported that persons who feel their family might be at risk would be less likely to report to work (Beaton et al. 2008).

Naturally, healthcare organizations must have a disaster plan; however, within the organization the plans may exist at different levels. Unfortunately, even the best planning does not make hospitals immune from significant problems. Following Hurricane Sandy in 2012, challenges were reported with infrastructure failures related to electricity and communication and with community collaboration over resources to include fuel, transportation, hospitals beds, and public shelters (HHS Office of Inspector General

2015). Likewise, Mercy Hospital in Joplin, Missouri, was unprepared for the total destruction of its facility during a 2011 tornado (Braun 2012). The bottom line from an organizational perspective on planning is that it is almost impossible to truly envision the worst case unless it is total destruction. Therefore, planning must be comprehensive and cannot presume continued electricity, operating generators, or other "niceties."

page 329

Community-level planning at the town, city, county, state, or national level relies heavily on the quality of the planning done at the lower levels, especially for hospitals and healthcare organizations. It is imperative that those doing the planning and carrying out the plan remember that the nature of disasters is dynamic and in flux, requiring constant adjustment and, ultimately, learning from previous disasters (Beaton et al. 2008).

Considerations for Disaster Planning

This section explores the different components of disaster planning, especially related to information system and information process continuity. The four generally accepted components are planning (or mitigation or prevention), preparedness, response, and recovery (Smith and Macdonald 2006; Beaton et al. 2008; Reynolds and Tamanaha 2010; Gibson et al. 2012). The first two components must be accomplished prior to the disaster, with the third constituting the response to the disaster, and the fourth including an analysis of how well the organization performed during the disaster. Each is examined in turn.

Planning

Sometimes called planning, mitigation, or prevention, the first step in the process is to involve all stakeholders so that understanding, agreement, and consensus can be reached regarding the goals for the disaster plan under development (Beaton et al. 2008). In the hospital this would include the executives, but also the frontline workers and key representatives from all business units or functional areas. A good place to start is with the goal of the disaster plan. The BIA is an excellent tool for understanding what is needed to achieve the goals. For example, continuing full operations with admissions still being accepted is quite a different goal than the orderly shutdown of operations. Neither is better or worse than the other, but will be dependent on the mission or business focus of the healthcare organization. Generally, acute-care hospitals with emergency departments must continue operations, whereas a clinic or specialty hospital might be able to discontinue clinical operations.

The second step for the planning process is to conduct a risk assessment or hazard vulnerability assessment (HVA; Herzig 2010). This involves identifying persons, situations, or events that pose a particular danger or are particularly vulnerable (Beaton et al. 2008). For example, newborns and the elderly are more vulnerable to environmental changes or extremes. Thus, it would be important to consider the needs of special populations such as these in the disaster plan.

These two steps are really the prep work for the disaster plan, ensuring that all of the necessary parties are involved and special situations that might change the plan are identified.

Preparedness

The next step is preparedness. This largely entails determining how the different risks that have been identified will be handled. Risks can be removed, reduced, assigned, or accepted (Nelson 2008). The preparedness component includes several facets.

page 330

1. Developing the plan
2. Organizing to carry out the plan
3. Training

4. Acquiring equipment
5. Testing or disaster drills
6. Evaluating
7. Improving

The development of the plan requires significant input from information technology or systems because the success of the healthcare organization's disaster plan will be dependent on communication during the disaster. The plan should be concise, well-written, and easy to read. It should use generic job titles instead of names, and define the command structure and responsibilities (Herzig 2010).

Recommendations that have been learned and shared by organizations that have survived other disasters should be included in the plan. First and foremost is the importance of communication. Specifically, many organizations will have a telephone tree in place for contacting employees in case of emergency. However, those who have tried to communicate via telephone have found that it often does not work. Fortunately, texting often does work because it requires much less bandwidth than maintaining an open line for a call (Harrison et al. 2008; Herzig 2010). In one disaster report, a flood severely displaced staff. The organization put information including emergency contact information on its website; however, this had not been a part of the communication plan so people did not know to check there (Anonymous 2011). Another important part of the plan organization is to ensure that roles and responsibilities are clear. Everyone should understand at least their initial tasks when disaster strikes (Beaton et al. 2008).

Training is a very essential part of preparedness. As shared in the previous paragraph, if staff are not trained for the ultimate communication method, it can be rendered ineffective. One reference maintains that all staff should know the following:

- Where the disaster plan is located
- Which staff members are essential and should be contacted
- Emergency contact details for all staff members
- Evacuation procedures for their department or work location
- The location of any disaster records in the hospital
- How to access any electronic lists of patient details that exist. (Smith and Macdonald 2006)

Equipment acquisition for preparedness can include everything. Examples of these considerations are determining where, when, and how a hot backup of the EHR will operate; ensuring that adequate emergency documents are available in selected storage sites; ensuring solar-powered batteries are purchased and available for cell phones or other communication devices if necessary; and ensuring adequate water and food are available for patients and essential staff in an emergency (Beaton et al. 2008). Because these decisions for additional equipment or supply acquisition could involve considerable expense, they are usually made at the executive or higher levels.

Healthcare facilities are required to conduct training, disaster drills, and other exercises to maintain their accreditation (Beaton et al. 2008). These exercises and drills often reveal weaknesses or other deficiencies that can then be addressed in the disaster plan. Any backup systems or remote storage or anything not used day-to-day should be tested to ensure it works as planned.

Ongoing evaluation of the disaster plan is essential. At a minimum the plan should be reviewed annually or more often if significant changes to operations or information systems are made. Changes, including new clinical processes or equipment, are easy to

implement without consideration given to disaster preparedness. Of course, the main reason for evaluation is to continue to improve the disaster plan and preparedness.

Although organizations must go through the planning and preparedness components, they do it in the hopes that they will never have to experience the next two phases, response and recovery.

Response and Recovery

Activation of the disaster plan can be chaotic due to the nature of the disaster that has occurred, but a comprehensive plan and adequate training will expedite the response. When a disaster actually occurs, it is not likely to unfold exactly as anticipated and planned for in a number of ways, such as intensity, destruction and collateral damages, disruption of services, and number and types of patients affected. For example, a plan for a tornado would include the potential for large numbers of injured to be treated at the hospital, but it is unknown who these people might be. The medical needs of those injured from an elementary school would be quite different than those of the injured from a long-term care facility. Such a situation would require a quick analysis of the circumstances to determine what adjustments are required to carry out the plan. Flexibility, ingenuity, and rapid decision-making are required to address the needs as presented. An example of adjusting to meet the needs in a disaster as they unfolded occurred after a 2013 tornado in Moore, Oklahoma. The storm arose quickly, allowing for little warning. It was classified as an EF-5 tornado up to two miles wide, with winds over 200 mph. More than 240,000 structures were damaged or destroyed, including two elementary schools and the only hospital in the community (NIST Engineering Laboratory 2014). Tornadoes are a known risk in this geographic area and significant planning and preparation is done to prepare for them; however, the intensity and type of destruction could not have been anticipated, nor could the fact that the hospital would be a part of the destruction. Other healthcare facilities in the area had to quickly be engaged to accommodate the number of injured, which included a large number of elementary school children, and a communication plan had to be established to communicate with their parents. Following the initial response phase immediately after the tornado, preparation began to provide long-term safety and appropriate care for the residents without a hospital in the immediate area. Having a disaster plan was important for this event to facilitate the response to such an occurrence, but that tornado was a prime example of when flexibility, ingenuity, and problem solving is vital when an actual disaster occurs.

Once past the response to a crisis, it is important to use recovery to learn from the experience. It is also important to pay attention to the psychosocial effects of a disaster on the workforce. Once the worst of the crisis passes, people can experience posttraumatic stress or depression as they learn how to cope with changed circumstances, sometimes in both their personal and their work lives.

Case Study: Ransomware

Hackers often penetrate a CE's network and get access to ePHI by exploiting known vulnerabilities. After they get access the ePHI, the hackers often encrypt the ePHI to hold it for ransom. This is called a ransomware attack. The Internet Crime Complaint Center at the Federal Bureau of Investigation reported that there were at least 649 ransomware attacks on critical infrastructure organizations between June and December 2021. Among all critical infrastructure sectors, the healthcare and public health

sector was the worst affected. There were 148 ransomware attacks in healthcare and public health sector in that period of time. Although hackers used many methods to access victim networks, the most common attack vectors in 2021 were phishing emails. The full implementation of the HIPAA Security Rule can help prevent the introduction of ransomware. OCR issues a guideline to help CEs and BAs better understand and respond to the threat of ransomware (OCR 2016b).

Using the knowledge you learn from this chapter, can you come up with some prevention actions to avoid ransomware attacks? You may check your list again the ransomware guide part 1: Ransomware prevention best practices created by the Cybersecurity & Infrastructure Security Agency: https://www.cisa.gov/stopransomware/ransomware-guide

Case Study: Telehealth

Telehealth allows your doctor to provide care to you when you cannot be in the doctor's office in person. It has been available for many years. However, most people still used to go to the doctor's office and receive care, even if the telehealth option was available to them. Less than one percent of healthcare visits were conducted via telehealth before 2020. Since 2020, the COVID-19 pandemic has had a dramatic effect on telehealth visits because of lockdowns, travel restrictions, social distance requirements, and masking policies. Almost all hospitals and doctor's offices started to offer telehealth visits to their patients. The CMS and insurance companies also dramatically expanded telehealth services to wider geographic locations and across state lines. This made more healthcare services available to patients during the pandemic. On the other hand, because most hospitals and doctor's offices had never offered telehealth services to their patients before the pandemic and they had to ramp up to provide the telehealth services in a very short period of time, some chose to use readily available but not HIPAA-compliant video conferencing systems to deliver their healthcare services. That was a risky action and a violation of the HIPAA Security Rule.

Further readings: Go to https://healthitsecurity.com/ and search for telehealth, and read some recent articles related to telehealth practices

Chapter 11 Review Exercises

Answer the following questions.

1. Which of the following is not considered to be identity theft?
   a. Using information accessed in a patient chart to receive medical care
   b. Information inadvertently put in the wrong patient's chart
   c. Medical care solicited using insurance information obtained from another's insurance card
   d. Using patient identification information to buy drugs
2. Red flags can include:
   a. Original unaltered documents (insurance card, driver's license)

b. Inconsistent address or other information on documents
c. Two forms of valid ID
d. Consistent birthdates on all documents

3. Which organization(s) issued the regulations for Red Flag Rules for healthcare?
   a. Federal Trade Commission
   b. Department of Treasury
   c. Federal Reserve System
   d. HHS
4. One method for evaluating and prioritizing risks related to the protection of organizational information is:
   a. Business impact analysis
   b. Recovery point objective
   c. Recovery time objective
   d. Disaster preparedness

5. Which of the following is not a required element of a security risk safeguard?
   a. The means of data collection for the organization
   b. Documentation of potential threats
   c. Assessment of the current security measures
   d. A contingency plan

Indicate whether the following statement is true or false

6. Other than the fact that it occurs in a healthcare facility, medical identity theft is to the same as financial identity theft.

Discussion

7. Discuss the process for a security risk analysis.
8. What would you consider in preparing a comprehensive disaster plan?
9. Identify one recent natural disaster and determine its impact to the healthcare system in the corresponding region.
10. Identify one recent ransomware attack on ePHI and determine the attack vectors used in the attack and the damage to the corresponding healthcare facility.

REFERENCES

45 CFR 160.18: Compliance date. 2013.
45 CFR 160.103: Definitions. 2013.
45 CFR 160.404: Amount of a civil money penalty. 2013.
45 CFR 164.302: Applicability. 2013.
45 CFR 164.304: Definitions. 2013.
45 CFR 164.306: Security standards: General rules.
45 CFR 164.306(d)(3): Security standards: General rules: Implementation specifications.
45 CFR 164.306(e): Security standards: General rules: Maintenance.
45 CFR 164.308(a): Administrative safeguards. 2013.
45 CFR 164.308(a)(1): Administrative safeguards. 2013.
45 CFR 164.308(a)(1)(i): Administrative safeguards: Standard: Security Management Process.
45 CFR 164.308(a)(1)(ii)(A): Administrative safeguards: Implementation Specification: Risk analysis. 2013.

45 CFR 164.308(a)(1)(ii)(A)-(B): Administrative safeguards: Implementation Specification: Risk Analysis (required), Implementation Specification: Risk Management (required).
45 CFR 164.308(a)(2): Administrative safeguards: Standard: Assigned security responsibility. 2013.
45 CFR 164.308(a)(3): Administrative safeguards. 2013.
45 CFR 164.308(a)(4): Administrative safeguards. 2013.

45 CFR 164.308(a)(5): Administrative safeguards. 2013.
45 CFR 164.308(a)(5)(i): Administrative safeguards: Standard: Security Awareness and Training.
45 CFR 164.308(a)(5)(ii)(A): Administrative safeguards: Implementation Specification: Security Reminders (addressable).
45 CFR 164.308(a)(6): Administrative safeguards. 2013.
45 CFR 164.308(a)(7): Administrative safeguards. 2013.
45 CFR 164.308(a)(8): Administrative safeguards: Standard: Evaluation. 2013.
45 CFR 164.308(b)(1): Administrative safeguards: Business associate contracts and other arrangements. 2013.
45 CFR 164.310: Physical safeguards. 2013.
45 CFR 164.310(a)(1): Physical safeguards: Standard: Facility access controls. 2013.
45 CFR 164.310(b): Physical safeguards: Standard: Workstation use. 2013.
45 CFR 164.310(c): Physical safeguards: Standard: Workstation security. 2013.
45 CFR 164.310(d)(1): Physical safeguards: Standard: Device and media controls. 2013.
45 CFR 164.312(a)(1): Technical safeguards: Standard: Access Control.2013
45 CFR 164.312(b): Technical safeguards: Standard: Audit controls. 2013.
45 CFR 164.312(c)(1): Technical safeguards: Standard: Integrity. 2013.
45 CFR 164.312(d): Technical safeguards: Standard: Person or entity authentication. 2013.
45 CFR 164.312(e)(1): Technical safeguards: Standard: Transmission security. 2013.
45 CFR 164.014: Organizational requirements. 2013.
45 CFR 164.016: Policies and procedures and documentation requirements. 2013.
45 CFR 164.402: Definitions. 2013.
AHIMA. 2013. " ***Analysis of modifications to the HIPAA Privacy, Security, Enforcement, and Breach Notification Rules under the Health Information Technology for Economic and Clinical Health Act and the Genetic Information Nondiscrimination Act: Other modifications to the HIPAA Rules.*** " http://library.ahima.org/xpedio/groups/public/documents/ahima/bok1_050067.pdf.
AHIMA. 2017. ***Pocket Glossary for Health Information Management and Technology***, 5th ed. Chicago: AHIMA Press.
AHIMA e-HIM Work Group on Medical Identity Theft. 2008. Mitigating medical identity. ***Journal of AHIMA*** 79(7):63–69.
American Recovery and Reinvestment Act of 2009. Public Law 111-5. https://www.gpo.gov/fdsys/pkg/PLAW-111publ5/content-detail.html.
Anonymous. 2011. All IRBs should prepare for possible disaster interruptions. ***IRB Advisor*** 11(11):109–111.

Beaton, R., E. Bridges, M. K. Salazar, M. W. Oberle, A. Stergachis, J. Thompson, and P. Butterfield. 2008. Ecological model of disaster management. ***AAOHN Journal***

56(11):471–478.
Braun, J. 2012. HIPAA rules stay in place during disaster—Address privacy and security issues in plans. ***Hospital Access Management*** (August 2):1–2.
Centers for Medicare and Medicaid Services (CMS). 2007a. "HIPAA Security Series—Security Standards: Administrative Safeguards." https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/administrative/securityrule/adminsafeguards.pdf .
Centers for Medicare and Medicaid Services (CMS). 2007b. "HIPAA Security Series—Security Standards: Physical Safeguards." https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/administrative/securityrule/physsafeguards.pdf .
Centers for Medicare and Medicaid Services (CMS). 2007c. "HIPAA Security Series—Security Standards: Technical Safeguards." https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/administrative/securityrule/techsafeguards.pdf .
Conn, J. 2012. Bracing for a crash. ***Modern Healthcare*** 42(20):32–33.
Department of Health and Human Services (HHS). 2016. ***HIPAA Enforcement. Health Information Privacy***. http://www.hhs.gov/hipaa/for-professionals/compliance-enforcement.
Department of Health and Human Services (HHS), HealthIT. 2015. ***Guide to Privacy and Security of Electronic Health Information. Washington, DC: HHS.***
Department of Health and Human Services (HHS), Office of Inspector General. 2015. ***The Response to Superstorm Sandy Highlights the Importance of Recovery Planning for Child Care Nationwide***. Washington, DC: HHS.
Department of Health and Human Services (HHS). 2020. "Enforcement results by year." https://www.hhs.gov/hipaa/for-professionals/compliance-enforcement/data/enforcement-results-by-year/index.html .
Department of Health and Human Services (HHS). 2021. Adjustment of civil monetary penalties for inflation and the annual civil monetary penalties inflation adjustment for 2021. ***Federal Register*** 86(217).
Federal Trade Commission (FTC). 2013. ***Fighting Identity Theft with the Red Flags Rule: A How-To Guide for Business.*** Washington, DC: FTC. https://www.ftc.gov/tips-advice/business-center/guidance/fighting-identity-theft-red-flags-rule-how-guide-business .
Gibson, P., F. Theadore, and J. Jellison. 2012. The common ground preparedness framework: A comprehensive description of public health emergency preparedness. ***American Journal of Public Health*** 102(4):633–642. doi:10.2105/AJPH.2011.300546.
Gupta, A. 2011. Hackers, breaches, and other threats to elect. ***Health Data Management*** 19(9):3.
Harrison, J. P., R. A. Harrison, and M. Smith. 2008. Role of information technology in disaster medical response. ***Health Care Manager*** 27(4):307–313.
Herzig, T. W., ed. 2010. ***Information Security in Healthcare: Managing Risk***. Chicago: HIMSS.
Keith, B. 2010. Catch fraud upfront: Medical identity theft is on the rise: It's time to revamp your processes. ***Hospital Access Management*** 29(5):49–51.
King, P., and R. L. Williams. 2008. ***Red Flag Compliance for Healthcare Providers: Protecting Ourselves and Our Patients from Identity Theft***. Chicago: American Health Lawyers Association and HIMSS.

National Institute of Standards and Technology (NIST). 2008. "Special Publication 800-66: A ***n introductory resource guide for implementing the Health Insurance Portability and Accountability Act (HIPAA) Security Rule."*** http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-66r1.pdf.

National Institute of Standards and Technology (NIST). 2016. "Publications Series 800-66: ***An introductory resource guide for implementing the Health Insurance Portability and Accountability Act.***" http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-66r1.pdf.

National Institute of Standards and Technology (NIST) Engineering Laboratory. 2014. "***Tornado, Moore, Oklahoma."*** http://www.nist.gov/el/disasterstudies/weather/moore_tornado_2013.cfm.

National Institute of Standards and Technology (NIST). 2022. NIST National Vulnerability Database. https://nvd.nist.gov.

Nelson, S. B. 2008. Information management during mass casualty events . . . Includes discussion. ***Respiratory Care*** 53(2):232–238.

Office for Civil Rights. Department of Health and Human Services (OCR). 2010. "***Guidance on risk analysis requirements under the HIPAA Security Rule.***" http://www.hhs.gov/ocr/privacy/hipaa/administrative/securityrule/rafinalguidancepdf.pdf.

Office for Civil Rights, Department of Health and Human Services (OCR). 2013. Modifications to the HIPAA Rules. ***Federal Register*** 78(17):5566–5702.

Office for Civil Rights, Department of Health and Human Services (OCR). 2016a. "***Health information privacy, Breach Notification Rule.***" http://www.hhs.gov/hipaa/for-professionals/breach-notification/index.html.

Office for Civil Rights, Department of Health and Human Services. 2016b. "***HIPAA security guidance: Ransomware.***" https://www.hhs.gov/sites/default/files/RansomwareFactSheet.pdf?language=es.

Office for Civil Rights. Department of Health and Human Services (OCR). 2021. ***Fall 2021 Cybersecurity Newsletter.*** https://www.hhs.gov/hipaa/for-professionals/security/guidance/cybersecurity-newsletter-fall-2021/index.html .

Office for Civil Rights, Department of Health and Human Services (OCR) 2022a. Breach portal. https://ocrportal.hhs.gov/ocr/breach/breach_report.jsf.

Office for Civil Rights, Department of Health and Human Services. (OCR). 2022b. ***Quarter 1 2022 Cybersecurity Newsletter***. https://www.hhs.gov/hipaa/for-professionals/security/guidance/cybersecurity-newsletter-first-quarter-2022/index.html .

Ranajee, N. 2012. Best practices in healthcare disaster recovery planning. ***Health Management Technology*** 33(5):22–24.

Reynolds, P., and I. Tamanaha. 2010. Disaster information specialist pilot project: NLM/DIMRC. ***Medical Reference Services Quarterly*** 29(4):394–404. doi:10.1080/02763869.2010.518929.

Smith, E., and R. Macdonald. 2006. Managing health information during disasters. ***Health Information Management Journal*** 35(2):8–13.

Springer, R. 2009. Medical identity theft—Red flag and address discrepancy requirements. ***Plastic Surgical Nursing*** 29(2):131–134. doi: 10.1097/01.PSN.0000356874.56338.2c.

ADDITIONAL RESOURCES

Amesh, A., M. Watson, N. Boeri, K. Minton, R. Morhard, and E. Toner. 2014. Absorbing citywide patient surge during Hurricane Sandy: A case study in accommodating multiple hospital evacuations. ***Annals of Emergency Medicine.*** https://www.jointcommission.org/assets/1/6/Hurricaine_Sandy_Hospital_Evacs_adalja2014.

Belfort, R. 2009. "***HITECH raises the stakes on HIPAA compliance.***" https://www.manatt.com/insights/newsletters/health-update/client-alert-%c2%a0hitech-raises-the-stakes-on-hipaa-coht .

Brocato, L., S. Emery, and J. McDavid. 2011. Keeping compliant: Managing rising risk in physician practices. ***Journal of AHIMA*** 82(11):32–35.

Cybersecurity and Infrastructure Security Agency (CISA), National Cyber Awareness System. 2022. "Alerts." https://us-cert.cisa.gov/ncas/alerts.

Cybersecurity and Infrastructure Security Agency (CISA), National Cyber Awareness System. 2022. "Bulletins." https://us-cert.cisa.gov/ncas/bulletins.

Dinh, A. 2008. HIPAA security compliance—What comes next? ***Journal of Health Care Compliance*** 10(3):37–39.

Health Information Technology for Economic and Clinical Health (HITECH) Act. Public Law 111-5.

Health IT Security, https://healthitsecurity.com/topic/healthcare-cyber-security.

HIPAA Journal, https://www.hipaajournal.com/.

Hoffman, S., and A. Podgurski. 2009. E-health hazards: Provider liability and electronic health record systems. ***Berkeley Technology Law Journal*** 24(4):1523–1582.

Neal, D. 2011. Choosing an electronic health records system: Professional liability considerations. ***Innovations in Clinical Neuroscience*** 8(6):43.

Rhinehart-Thompson, L. A. 2008. Raising awareness of medical identity theft: For consumers, prevention starts with guarding, monitoring health information. ***Journal of AHIMA*** 79(10):74–75, 81.

Schiff, G. D., and D. W. Bates. 2010. Can electronic clinical documentation help prevent diagnostic errors? ***New England Journal of Medicine*** 362(12):1066–1069. doi: 10.1056/NEJMp0911734.

Scott, R. L. 2006. ***Physicians' Reliance on Electronic Health Records.*** Houston, TX: University of Houston.

Sullivan, J. 2004. ***HIPAA: A Practical Guide to the Privacy and Security of Health Data***. Chicago: American Bar Association.

Vigoda, M. 2008. e-Record, e-Liability: Addressing medico-legal issues in electronic records. ***Journal of AHIMA*** 79(10):48–52.

Virapongse, A., D. W. Bates, P. Shi, C. A. Jenter, L. A. Volk, K. Kleinman, L. Sato, and S. R. Simon. 2008. Electronic health records and malpractice claims in office practice. ***Archives of Internal Medicine*** 168(21):2362–2367. doi:10.1001/archinte.168.21.2362.

Verizon. 2020. "Data breach investigations report." https://enterprise.verizon.com/resources/reports/2020/2020-data-breach-investigations-report.pdf.

Chapter

12

Data Governance

By Lesley Clack, ScD, CPH

Learning Objectives

- Explain data governance and its importance to the field of health informatics
- Analyze the key drivers for managing information in an organization
- Explain the process for effective data governance in an organization
- Define data quality and explain why data quality is important in data governance
- Differentiate between data governance and information governance

KEY TERMS

Accessibility of data
Accuracy of data
Analysis of data
Application of data
Big data
Collection of data
Comprehensiveness of data
Consistency of data
Currency of data
Data
Data governance
Data management
Data quality
Definition of data
Granularity of data
Information governance
Precision of data

Relevance of data
Timeliness of data
Value of data
Variety of data
Velocity of data
Veracity of data
Volume of data
Warehousing of data

Healthcare data are abundant and are retrieved from multiple sources. In healthcare, data often come from patient observations or patient input (Amatayakul 2017). Healthcare data are complex, are both structured and unstructured, and come from a variety of sources. Thus, it is necessary to have processes in place for standardizing the data for sharing both within and outside the organization.

Data Management

Data are an asset to an organization, similar to the its personnel, buildings, and other resources. As such, data require attention and maintenance to ensure they are capable of supporting organizational needs and goals. The management of data is as important to a healthcare organization as the management of its buildings, capital equipment, and information technology (IT) infrastructure. **Data management** has been defined as "the practice of collecting, keeping, and using data securely, efficiently, and cost-effectively" (Oracle 2022). The process of data management involves everything from creating, accessing, updating, and storing and using data to ensuring data privacy and security and proper archiving and destruction of data. Data management is an overarching process that involves oversight of many areas, such as managing databases and configuring systems to make sure that accurate information is produced (Amatayakul 2017). Management of digital data includes creating and updating data, storing data, and using data on a variety of different platforms (Oracle 2022).

Data Governance

Data governance is an essential component of the management of data in the field of health informatics and is an emerging practice in the healthcare industry. AHIMA defines **data governance** as "the overall administration, through clearly defined procedures and plans, that assures the availability, integrity, security, and usability of the structured and unstructured data available to an organization" (AHIMA 2020, 1023). Data governance has also been defined as "the set of policies and procedures that determine the who, how, and why of data management within the organization" (Nunn 2009, 50).

Data governance and data management are two separate concepts, and it is important to understand the difference. Data management involves making and implementing decisions, whereas data governance is concerned with what decisions should be made to ensure the effective management of the data (Khatri and Brown 2010). Data governance happens from the top down in an organization, it is strategic, and it answers what needs to happen. Data management happens from the bottom up, is tactical, and answers how to get it done (Empel 2014). Data governance arrived as a concept in the field of health information management (HIM) after transaction systems and coding technologies led to the availability of structured data with high integrity (AHIMA 2011). Key functions for

data governance include (1) approving data management plans, policies, processes and standards; (2) assigning accountability and responsibility for owners, stewards, and custodians of physical data storage; (3) developing decision rights and change processes; (4) addressing regulatory and other external requirements; and (5) enforcing compliance (ONC 2022).

Figure 12.1. Key drivers for effectively managing information in an organization

Cost Retention of storage of information is costly to the organization, so HIM professionals need to assist the organization by helping to identify what information needs to be retained in accordance with applicable laws, regulations, and policies.
Meaningful Use There may be regulatory incentive programs that require maintaining and using information in a meaningful way, so organizations must ensure that these laws, regulations, and policies are followed.
Innovation Due to the diversity of sources of information in an organization, this can lead to innovative models being created for accessing the diverse information sources.
Patient Safety Protecting information and ensuring accurate and complete data supporting patient safety improvement efforts.
Healthcare Customers Healthcare customers want access to their information in a timely and efficient manner, which aids in making informed decisions about their care.
Legal/Compliance Pressures While retention of information is costly, organizations must balance the expense with compliance of laws and regulations that require certain standards are followed in regard to the management of information in an organization.

Source: AHIMA 2011.

The key drivers that push organizations to more effectively manage their information are shown in figure 12.1. These drivers are crucial for organizations in making decisions on how to most efficiently and effectively manage data.

Data Governance Process

As with most important healthcare processes, data governance is not a one-time effort, nor should it be "fast-tracked". It is important to take a careful approach, as the decisions made for data governance will have a long-term impact on the organization. A suggested process follows:

1. Discover—Identify the databases or data sets or applications that store data, the relationships between the different data sets, the meaning of the data to the organization, and who has responsibility for the data. This step requires the business users to be involved.

2. Design—Consolidation and coordination of organizational data and the environment focusing on consistency of any governing rules, consistency of the organizational data model, and consistency of the business processes. This step requires the involvement of the business users.
3. Enable—This step takes a considerable amount of time and effort, as this is when the new data governance standards are applied to each data source, business process, and application. It is recommended that the standards be deployed to the entire network, rather than embedded in each source, process, or application. Although this includes business users, much of the responsibility falls on IT staff.
4. Maintain—A robust monitoring and reporting system is required to ensure that the data remain fit-for-purpose for the organization. The monitoring requires involvement from both business users and IT. The monitoring should be as automated as possible, with reporting to the responsible parties, who then have the authority to make changes resulting in optimal standards and processes. Any changes must be documented so the data can be thoroughly understood when they are used.
5. Archive or Retire—It is important to ensure that the data that are no longer needed are retired in a methodical, considered fashion. This may involve complying with the legal requirements for maintaining the data, as well as understanding the different uses of the data within the organization. (Fisher 2012).

Table 12.1 . Data governance framework

Domain	Description
Data principles	An organization's data principles take into account the usage of external data and the regulatory environment that could influence the business uses of data. Data principles are discussed in further detail later in this chapter.
Data quality	Data quality is essential because poor data quality can affect an organization at both operational and strategic levels. Data quality is discussed in further detail later in this chapter.
Metadata	Metadata describes what the data are about, which assists in the interpretation of the data. Metadata are discussed in further detail in Chapter 8.
Data access	Data access refers to the ability to access or retrieve data within a database.
Data life cycle	Data life cycle refers to the fact that all data moves through life-cycle stages.

Source: Khatri and Brown, 2010.

Data governance and stewardship are relatively new aspects of data management in healthcare organizations. Although the processes can be difficult and time consuming, they are essential to any organization wishing to maximize the use of its valuable data assets.

Data Governance Framework

A framework for data governance was developed by Khatri and Brown (2010). The framework includes five interrelated decision domains, as are shown in table 12.1.

Check Your Understanding 12.1

Answer the following questions

1. Regulatory incentive programs that require maintaining and using information in a meaningful way so that organizations must ensure that these laws, regulations, and policies are followed are called:
   a. Data Principles
   b. Data Security Measures
   c. Meaningful Use
   d. Metadata
2. _____ takes into account the usage of external data and the regulatory environment that could influence the business uses of data.
   a. Data access
   b. Data life cycle
   c. Data principles
   d. Data quality
3. Which of the following describes what the data are about, which assists in the interpretation of the data?
   a. Data access
   b. Data life cycle
   c. Data principles
   d. Metadata
4. The ability to retrieve data within a database is called:
   a. Data access
   b. Data life cycle
   c. Meaningful Use
   d. Metadata
5. Which of the following refers to the stages that data move through?
   a. Data access
   b. Data life cycle
   c. Data principles
   d. Data quality

Discussion.

6. Explain the difference between data management and data governance.
7. Of the key drivers for effectively managing information in an organization, which one do you think is the most important and why? Explain your answer.

Data Quality

Assessing and understanding the quality of healthcare data form an extremely important and often overlooked aspect of managing data and information in healthcare provider settings. Ensuring that the data and information are correct has implications for the quality of patient care, reimbursement, and research. In short, it is important for all of the reasons that providers document in patient records to begin with. For example, something as simple as the wrong patient weight could result in an incorrect drug dosage calculation and medication administration error. Rather than getting bogged down in the details of studies regarding data quality, this section discusses various aspects of data quality and draws from more general computer information systems theory to suggest methods for assuring high-quality data and information for healthcare.

Characteristics of Data Quality

Data quality can seem an overwhelming topic. There are many ways for data to be of poor quality and, in fact, the quality of the data can sometimes resemble the quality of patient care—it can vary according to the consumer. AHIMA defines **data quality** and integrity as "the extent to which healthcare data are complete, accurate, consistent, and timely throughout its lifecycle including collection, application (including aggregation), warehousing, and analysis" (AHIMA 2022, 2). "Data are of high quality if they are fit for their intended uses in operations, decision making, and planning. Data are fit for use if they are free of defects and possess desired features" (Redman 2006, 73). This emphasizes that the data consumer and the use of the data are essential to any determination of data quality. In 1998, the AHIMA Data Quality Management Task Force developed a Data Quality Management Model, most recently updated in 2015 (Davoudi et al. 2015). The current model is shown in figure 12.2. It covers the different processes of data handling during which data quality should be addressed. The data handling processes are as follows:

- **Application of data**—The purpose for which the data are collected
- **Collection of data**—The processes by which data elements are accumulated
- **Warehousing of data**—Processes and systems used to archive data and data journals
- **Analysis of data**—The process of translating data into information utilized for an application (Davoudi et al. 2015).

In addition, 10 characteristics of data quality were identified and defined. The definitions and examples are provided in figure 12.2. The 1998 model was updated in 2015. Although the fundamentals remain the same, understanding of how they can most effectively be implemented in healthcare continue to evolve.

Figure 12.2. Data quality management model

▶ Expand/collapse Extended Description

Source: AHIMA 2015.

It is important to understand that these processes and characteristics do not exist in a vacuum. Just as everyone who has a driver's license has to meet certain criteria, but not all drivers are equally accomplished, all data that meet these criteria might still not be appropriate or adequate. This is why the process of assessing data quality is so significant. Failure to address the quality of the data could mean that the massive increase in data collected and available due to the increased use of computers in healthcare has only resulted in an increase in poor quality data and, hence, inaccurate decision-making.

The Data Quality Assessment and Management Process

The process of data quality assessment is similar to and, in some instances, uses the same methods as healthcare quality assessment and management. This section outlines generally accepted steps to quality assessment through the lens of quality management. It has been explicated using examples from the healthcare industry.

Who Are the Data Consumers?

Any health informatics professional tasked with assessing or managing data quality must first understand who the data consumers are. This involves making a list of all of the internal and external data consumers. Likely internal data consumers for healthcare data managers include patients, clinicians, administrators, and researchers. External data

consumers might include payers, public health agencies, and law enforcement agencies. Once a complete list has been compiled, the consumer list must be prioritized. This is an essential step. It will be impossible for the data manager to meet all of the needs of all of the data consumers. However, by enumerating all of the data consumers, data managers will begin documenting the full depth and breadth of data use in their organization. page 348

Table 12.2 . AHIMA data quality characteristics, definitions, and examples

Data quality characteristic and definition	Healthcare example
Accuracy of data—The extent to which the data are free of identifiable errors.	The data element gender is completed for all patients and a random check of 500 records performed annually revealed only one demographic data element in conflict with the documentation.
Accessibility of data—Data items are easily obtainable and legal to access with strong protections and controls built into the process.	All persons with access to the EHR have the ability to search the master patient index and the search function is designed such that the wrong patient is rarely accessed.
Comprehensiveness of data—All required data items are included; ensures that the entire scope of the data is collected with intentional limitations documented.	Providers may decide that recording external cause data is not useful. If so, their data dictionary for the diagnoses data elements would need to include this as an intentional limitation.
Consistency of data—The extent to which the healthcare data are reliable and the same across applications.	Within the EHR, data such as allergies must be consistently displayed within different applications or screens to prevent confusion.
Currency of data—The extent to which data are up to date; a datum value is up to date if it is current for a specific point in time, and it is outdated if it was current at a preceding time but incorrect at a later time.	Patient date of birth is generally current when the care is delivered. If date of birth is not re-entered, the software needs to have functionality to automatically update the age when appropriate.
Definition of data—The specific meaning of a healthcare-related data element.	Address as a data element label can mean the street address or it can mean the entire address to include city, state, and zip code.
Granularity of data—The level of detail at which the attributes and values of healthcare data are defined.	Adult weights are usually only recorded in pounds, possibly tenths of a pound. Newborn weights must be recorded in terms of ounces for accuracy.
Precision of data—Data values should be strictly stated to support the purpose.	Diagnosis-related group (DRG) values are carried out to four digits behind the decimal. It would be inaccurate to have the system only use two digits behind the decimal.
Relevance of data—The extent to which healthcare-related data are useful for the purposes for which they were collected.	Recording a primary diagnosis for hospital inpatients would be irrelevant because coding guidelines mandate collection of the principal diagnosis, which can be entirely different.

page 349

Timeliness of data—Concept of data quality that involves whether the data are up-to-date and available within a useful time frame; timeliness is determined by the manner and context in which the data are being used.	It would be inappropriate for blood pressure readings from ICU monitors to only be updated each hour.

Source: Davoudi et al. 2015

What Are the Needs of the Data Consumers?

This step in the process is usually very challenging. Most data consumers do not understand their needs well enough to articulate them effectively. So, what is the data manager to do? Rather than asking what the data consumers need, the data manager should ask the consumers to describe how they use the data, the decisions that are made, the potential consequences of a wrong decision, and so on. As an example, consider the data use of clinical decision support. Even within this seemingly single data use task there are nuances. Allergic drug reactions as well as preventive service reminders are each considered clinical decision support. However, allergic drug reactions are much more important because incorrect data could result in immediate, serious patient injury, or even death. Although preventive service reminders are certainly important, they rarely, if ever, rise to the level of immediate injury or death. The person working with the consumers to determine their needs must explore each data use thoroughly.

The data manager must take a controlled and careful approach to collecting the data that the data consumer needs. The initial steps of implementing a data quality process will not be quick. However, the data manager should better understand the needs of the various consumers after completion of the template or similar tool. In addition, the data manager will now have a complete list of data consumer needs to carry forward for consideration by senior management if necessary. It would be usual and unexpected for senior management to become involved in data consumer needs prioritization efforts. Figure 12.3 is an example of a needs and requirements collection template that might be used.

Consumer:																									
Consumer needs						Product feature requirements									Gap analysis										
Primary Needs	Secondary Needs	Tertiary Needs	Translated Needs	Source	Unit of Measure	Accuracy	Accessibility	Comprehensiveness	Consistency	Currency	Definition	Granularity	Precision	Relevancy	Timeliness	Accuracy	Accessibility	Comprehensiveness	Consistency	Currency	Definition	Granularity	Precision	Relevancy	Timeliness

Figure 12.3. Data consumer needs assessment tool

Source: Adapted from Redman 2001

What Are the Required Features and Quality Characteristics?

After documenting the consumer needs, the functionality requirement associated with each quality characteristic needs to be identified. Building on the drug allergy reaction alert data use, each data quality characteristic will be examined in light of the functionality required. The following might be a partial list of the requirements:

- Accuracy—The lists of drugs will need to utilize the current EHR data standard for pharmaceuticals. Allergies will need to be recorded as structured data or processed using an NLP algorithm after entry to be usable for drug allergy reaction processing.
- Accessibility—The data entry methods for the allergies and drugs will be dependent on the workflow of the organization and the abilities of the persons entering the data. This requires further study.
- Comprehensiveness—Allergies are a required data element for the EHR. If the patient has no allergies the field should reflect "No allergies reported." If the clinician is not able to determine whether the patient has allergies the field should reflect "Unknown."
- Consistency—Once recorded, an allergy should become a permanent entry in the patient EHR, which is always incorporated into the information presented to the clinician.

Once the functionality requirements are identified across all of the characteristics, the data consumer should review them to ensure they meet consumer needs. Seeing the requirements specified in this fashion could result in the consumer identifying additional, important uses that will also need to be specified using these methods.

The end result of the requirements and quality characteristics specification involves a dual examination of both the stringency of the requirements and a prioritization of all requirements. The most stringent requirement might not be at the top of the prioritization list, but it is the highest level of functionality required.

How Well Do Our Current Information Products Meet the Needs and Requirements?

After identification of the most important needs and requirements, it is now time to determine how well any current information products perform. There are several steps for this process:

1. Ask the data consumers.
2. Measure the data where possible and appropriate.
    a. Select a business process or operation that uses the suspect data.
    b. Determine a desired sample size and selection method.
    c. Decide on the assessment criteria. In healthcare we would like to strive for 100 percent accuracy or correctness; however, there may be instances in nonclinical data where data are "seriously flawed" or "flawed but acceptable."
    d. Establish the estimated impact by asking the people who work with the end data. Do flawed data affect patient care? Does it affect the ability to seek reimbursement? What are the impacts on time, expenses, customer satisfaction, or other issues?
    e. Measure associated activities such as productivity and consumer satisfaction.
3. Include the results of the data measurement on the needs and requirements spreadsheet.

It is important to give in-depth consideration to documenting the data consumers' needs and concerns. Failure to do this can result in a product that does not meet the consumers' needs. Additionally, the consumers' representative has to sign off on the delivery of the final system, which must meet their functional specifications. Therefore, it is important to work closely with the data consumers to help them fully document the specifications of the new system. Another important consideration is how that functionality will be assessed. Consumers and information technologists are often challenged when designing tests of functionality. Additionally, consumers might be unfamiliar with the new technology that calls for integration into existing systems, which presents an additional layer of complexity for end user acceptance testing.

Where Are the Gaps and How Important Are They?

The data manager should now create a list of all of the gaps between the consumer needs and requirements and the current information products and performance identified in the previous steps. Once the gaps have been enumerated, the data manager can move forward with creating a priority list of areas or functionality needing attention. As always, this should not be done in a vacuum. The development of the list should be iterative, undergoing review by affected data consumers as well as senior management.

Check Your Understanding 12.2

Answer the following questions.

1. _____ refers to the extent to which the data are free of identifiable errors.
   a. Accessibility of data
   b. Accuracy of data
   c. Comprehensiveness of data
   d. Consistency of data
2. _____ refers to the extent to which data are reliable and the same across applications.
   a. Granularity of data
   b. Precision of data
   c. Comprehensiveness of data
   d. Consistency of data
3. Diagnosis-related group values are carried out to four digits behind the decimal. It would be inaccurate to have the system only use two digits behind the decimal. This refers to which characteristic of data quality?
   a. Precision of data
   b. Granularity of data
   c. Relevance of data
   d. Comprehensiveness of data
4. The extent to which data items are easily obtainable and legal to access refers to which characteristic of data quality?
   a. Currency of data
   b. Accessibility of data
   c. Accuracy of data
   d. Consistency of data

5. The level of detail at which the attributes and values of healthcare data are defined refers to which characteristic of data quality?
   a. Precision of data
   b. Granularity of data
   c. Relevance of data
   d. Comprehensiveness of data
6. It would be inappropriate for blood pressure readings from ICU monitors to only be updated each hour. This would affect which characteristic of data quality?
   a. Relevance of data
   b. Precision of data
   c. Timeliness of data
   d. Granularity of data

Case Study: Hospital Data Quality Management

A hospital may handle nearly 300,000 outpatients and 7,000 to 8,000 inpatients per year. As they frequently run incremental updates to their data warehouse to track orders and charges for each and every patient, they are tasked with processing and managing a vast amount of data complexity.

Prior to setting up a robust data quality management system, the hospital struggled to keep track of its numerous data transformation job completions and failures. By setting up proper monitoring systems, the staff is now able to determine exactly where such jobs fail. According to the data warehouse administrator at the hospital, they are now able to configure data integration jobs so that if there is a problem or exception somewhere in the data processing chain, email alerts or other error conditions are created.

In addition to monitoring their transformations, the hospital has set up systems to allow executives, physicians, and other clinicians to easily extract data and perform any transformation in preparation for reporting. By effectively allowing for collaboration and combining heterogeneous systems, the hospital is now able to effectively identify uniqueness across different systems and normalize the data into aggregated secondary data marts for validation and reporting. This has allowed for analyses that otherwise would have been unimaginable given the combinatorial complexity of these large systems.

Data Governance Management

According to the Office of the National Coordinator for Health Information Technology (ONC), the purpose of data governance management is to "establish the ownership, stewardship, and operational structures needed to ensure that data is managed as a critical asset and implemented in an effective and sustainable manner" (ONC 2022).

One of the principal challenges of information management is the governance of big data. **Big data** refers to a large collection of data sets that is complex and difficult to maintain and process (Joshi et al. 2019). Big data are defined as having more than one of the following, known as the Five Vs of Data:

1. **Volume of data**: An unusually large quantity of data.
2. **Variety of data**: A large range of different types of data that brings new data structures, standardizations, and data terminology.
3. **Velocity of data**: Data sources that are dynamic and have a high frequency of refreshes; for example, real-time data feeds.
4. **Veracity of data**: Data sources that are complex due to varying accuracy or completeness of the data.
5. **Value of data**: Data sources that vary in usefulness in terms of cost, workflow complexity, feasibility, and governance (Joshi et al. 2019).

Data Governance Structure

Establishing a data governance structure is the foundation for successful data governance (Ott 2015). Most governance structures have three layers: strategic, tactical, and operational (Gupta and Cannon 2020). The strategic layer approves any initiatives and funding requests and endorses outcomes and recommendations. Individuals who participate in the strategic level of data governance will assist in setting strategic direction for the organization, allocating resources, and approving policies and processes. The tactical layer is responsible for implementing the strategies and making recommendations for solutions to identified issues. Individuals who participate in the tactical level of data governance will work together to implement strategies, coordinate work, and suggest

solutions. The operational layer is responsible for implementation of suggested processes in accordance with policies and procedures. Individuals involved in the operational level of data governance will perform data management tasks in accordance with data governance policies and procedures (Gupta and Cannon 2020).

Ott proposed a 10-step plan for effective data governance structure (Ott 2015), shown in table 12.3.

Data Governance Versus Information Governance

An organization's data governance strategy should be integrated with its information governance strategy. The terms data governance and information governance are often used interchangeably, but they are two distinctly different concepts (Washington 2015). Information governance has been defined as "the specification of decision rights and an accountability framework to ensure appropriate behavior in the valuation, creation, storage, use archiving, and deletion of information" (Warner 2013). Information governance also refers to "the processes roles and polices, standards, and metrics that ensure the effective and efficient use of information in enabling an organization to achieve its goals" (Warner 2013).

Table 12.3 . 10-step plan for effective data governance structure

Step	Description
Step 1: Determine the strategy	To begin, the leaders can write a data governance charter, which should include a mission statement and the goals of the data governance program for the organization.
Step 2: Choose a model for a data governance team	The leaders should select the data governance model that is the best fit for the organization.
Step 3: Choose the right hierarchy for the organization	The leaders should select the appropriate hierarchy to demonstrate the role of data governance in the organization.
Step 4: Select the steering committee	The steering committee typically includes high-level executives of the organization and stakeholders who should direct the data governance process.
Step 5: Set up the data governance office	The data governance office enforces data governance in the organization.
Step 6: Choose the data governance working group	The data governance working group typically includes manager-level or higher leaders in the organization who are liaisons between business and IT.
Step 7: Select the data governance support team	The data governance support team should include the data owner, data steward, data architect, data modeler, and data analyst.
Step 8: Policy and procedure development and enforcement	The individuals in the data governance working group should create policies based on the program goals in the charter.
Step 9: Establish a skilled business team	Data governance teams should be comprised of both business experts and IT experts.
Step 10: Establish a skilled IT team	The IT team should be composed of individuals who are highly skilled in technology.

Source: Ott 2015

AHIMA developed Information Governance Principles for Healthcare (IGPHC), which establish a foundation of best practices for information governance programs in the form of eight principles (Empel 2014). These principles are shown in figure 12.4.

Figure 12.4. AHIMA Information Governance Principles for Healthcare (IGPHC)

AHIMA IGPHC Principles
1. Principle of Accountability Leaders of the organization are accountable for the information governance and should provide oversight and delegate responsibility to qualified individuals. Policies and procedures should be adopted to govern the organization's information governance.
2. Principle of Transparency The organization should be open and transparent about their processes and activities related to information governance. Documentation should be available to relevant stakeholders in accordance with the organization's policies and procedures on information governance.
3. Principle of Integrity Integrity of the information governance program should be ensured to protect information from breach, corruption, and loss, and to ensure that the privacy and confidentiality of information are protected.
4. Principle of Protection Information in an organization should be protected at the appropriate level from breach, corruption, and loss to ensure the privacy and confidentiality of the information.
5. Principle of Compliance Compliance with applicable laws, regulations, standards, and organizational policies and procedures is important for the information governance program.
6. Principle of Availability Information in an organization should be made available in a timely manner, and the organization should establish policies to ensure accurate and efficient retrieval of information.
7. Principle of Retention The organization should ensure that information is retained for the appropriate time in accordance with policies, laws, and regulations.
8. Principle of Disposition Once information is no longer required to be maintained, organizations should follow applicable policies, laws, and regulations to ensure appropriate disposition of information.

Source: Empel 2014

The Information Governance Adoption Model (IGAM) was created by AHIMA to introduce the constructs of information governance organizational core competencies that are further driven by performance markers. IGAM builds on the foundation of the IGPHC to define the characteristics of information governance programs at different levels of

maturity (Downing 2015). The information governance levels of maturity are shown in figure 12.5.

page 357

Figure 12.5. Information governance levels of maturity

Level	Description
Level 1 (At-Risk)	Level 1 is the level at which information governance concerns, requirements, and opportunities are not addressed at all, are addressed minimally, or are addressed in an ad hoc manner.
Level 2 (Aware)	Level 2 describes an environment in which awareness starts to develop surrounding the idea that information governance has an impact on the organization and that the organizations may benefit from a more defined information governance program.
Level 3 (Aspirational)	Level 3 describes an environment in which the organization has set forth policies and procedures and has implemented processes specifically intended to improve information governance.
Level 4 (Aligned)	Level 4 describes an environment in which the organization has established a proactive information governance program throughout its operations and has established continuous improvement for it.
Level 5 (Actualized)	Level 5 describes an environment in which the organization has integrated information governance into its overall infrastructure and business processes so that compliance with legal and regulatory requirements are routine.

Source: Downing 2015

Case Study: Data Governance Versus Information Governance

As discussed previously, data governance and information governance are two different concepts. Data governance is concerned with individual pieces of data, whereas information governance is concerned with the process and systems that use data for clinical decision-making. For example, a patient's cholesterol levels for the past year would fall under data governance. Information governance comes into play when the clinical provider aggregates the past year of cholesterol readings to diagnose the patient with hypercholesterolemia and prescribe specific medications.

page 358

Chapter 12 Review Exercises

Answer the following questions.

1. This principle refers to the ability of the information governance program to ensure information is protected from breach and corruption.
   a. Principle of transparency
   b. Principle of accountability
   c. Principle of integrity
   d. Principle of protection
   e. Both a and c
   f. Both b and c

g. Both c and d

2. Data that are dynamic and constantly changing refers to which of the Five Vs of Data?
   a. Variety
   b. Velocity
   c. Volume
   d. Veracity
3. Data that are complex due to varying accuracy or completeness of the data refers to which of the Five Vs of Data?
   a. Variety
   b. Velocity
   c. Volume
   d. Veracity
4. Ensuring that information meets all applicable laws, regulations, standards, and organizational policies and procedures refers to which information governance principle?
   a. Principle of accountability
   b. Principle of compliance
   c. Principle of protection
   d. Principle of disposition

Discussion

5. Explain the difference between data governance and information governance.
6. Why is data quality so important to data governance? Explain.
7. Of the information governance principles for healthcare, which do you think is the most important? Why?
8. Assume you are the director of the Health Informatics department in a healthcare organization. Explain how you would establish a data governance structure.

REFERENCES

AHIMA. 2011. HIM functions in healthcare quality and patient safety: Appendix B. HIM's role in data governance. ***Journal of AHIMA*** 82(8):expanded online version.

AHIMA. 2015. Data Quality Mangement Model (2015 Update)- Retired. https://library.ahima.org/PB/DataQualityModel#.YpYv-ajMKUk

AHIMA. 2020. ***Health Information Management, Concepts, Principles and Practice***, 6th ed. Chicago: AHIMA Press.

AHIMA. 2022. "Healthcare data governance." https://www.ahima.org/media/pmcb0fr5/healthcare-data-governance-practice-brief-final.pdf.

Amatayakul, M. K. 2017. ***Health IT and EHRs: Principles and Practice,*** 6th ed. Chicago: AHIMA Press.

Davoudi, S., J. A. Dooling, B. Glondys, T. D. Jones, L. Kadlec, S. M. Overgaard, K. Ruben, and A. Wendicke. 2015. Data quality management model (2015 Update)—Retired. ***Journal of AHIMA*** 86(10):expanded web version.

Downing, K. 2015. "AHIMA's IG maturity and adoption model turns the corner." AHIMA. https://bok.ahima.org/doc?oid=301679#.Ym33hdrMKUk.

Empel, S. 2014. Way forward: AHIMA develops information governance principles to lead healthcare toward better data management. ***Journal of AHIMA*** 85(10):30–32.

Fisher, T. 2012. ***The Data Asset: How Smart Companies Govern Their Data for Business Success.*** Hoboken, NJ: Wiley.

Gupta, U., and S. Cannon. 2020. ***A Practitioner's Guide to Data Governance: A Case-Based Approach.*** West Yorkshire, UK: Emerald.

Joshi, A., L. Thorpe, and L. Waldron. 2019. ***Population Health Informatics: Driving Evidence-Based Solutions into Practice.*** Burlington, MA: Jones & Bartlett.

Khatri, V., and C. V. Brown. 2010. Designing data governance. ***Communications of the ACM*** 53(1):148–152.

Nunn, S. L. 2009. Driving compliance through data governance. ***Journal of AHIMA*** 80(3):50–51.

Office of the National Coordinator for Health Information Technology (ONC). 2022. ***Health IT Playbook: Data Governance.*** https://www.healthit.gov/playbook/ambulatory-guide/data-governance/.

Oracle. 2022. "What is data management?" https://www.oracle.com/database/what-is-data-management/.

Ott, M. 2015. A ten-step plan for an effective data governance structure. ***The Data Administration Newsletter.*** https://tdan.com/a-ten-step-plan-for-an-effective-data-governance-structure/19183.

Redman, B. 2006. ***The Practice of Patient Education,*** 10th ed. St. Louis, MO: Mosby.

Warner, D. 2013. IG 101: What is information governance. ***Journal of AHIMA.***

Washington, L. 2015. "Information governance provides the framework for data governance." ***Journal of AHIMA.***

ADDITIONAL RESOURCES

Bhansali, N. 2014. ***Data Governance: Creating Value from Information Assets.*** Boca Raton, FL: Auerbach.

Dai, W., I. Wardlaw, Y. Cui, K. Mehdi, Y. Li, and J. Long. 2016. Data Profiling Technology of Data Governance Regarding Big Data: Review and Rethinking. In ***Information Technology: New Generations. Advances in Intelligent Systems and Computing***, edited by S. Latifi. Cham, Switzerland: Springer.

Fenton, S. H., and S. Biedermann. 2014. ***Introduction to Healthcare Informatics.*** Chicago: AHIMA Press.

Panian, Z. 2010. Some practical experiences in data governance. ***World Academy of Science, Engineering and Technology*** 62(1):939–946.

Petters, J. 2020. "Data governance in healthcare: Your complete guide." https://www.varonis.com/blog/data-governance-in-healthcare#:~:text=For%20example%2C%20a%20patient%E2%80%99s%20blood%20pressure%20readings%20for,patient%20with%20hypertension%20and%20advises%20a%20specific%20medication .

Talend. 2022. "The value of data quality in healthcare." https://www.talend.com/resources/data-quality-healthcare/#:~:text=Example%3A%20Data%20quality%20management%20at%20Omaha%20Children%E2%80%99s%20Hospital,outpatients%20and%207%2C000%20to%208%2C000%20inpatients%20per%20year .

Chapter

13

Ethics

By Lesley Clack, ScD, CPH

Learning Objectives

- Define ethics
- Understand basic ethical terms and concepts
- Explain the importance of ethical issues related to technology
- Analyze ethical issues related to health informatics
- Assess ethical standards of practice
- Apply the AHIMA Code of Ethics to ethical scenarios
- Apply the steps of the ethical decision-making process to a given scenario
- Develop policies and procedures for ethics related to informatics

KEY TERMS

Autonomy
Beneficence
Code of ethics
Confidentiality
Electronic health record (HER)
Ethical decision-making
Ethics
Hacker ethics
Health information exchange
Justice
Morals
Nonmaleficence
Personal health record (PHR)

Privacy
Release of information (ROI)
Secondary release of information
Security
Sentinel event
Social media
Social networking
Values
Whistleblower

Decisions of an ethical nature must be made on a daily basis when providing quality healthcare services across an organization and with individuals carrying out their professional responsibilities. Determining the course of action in these situations becomes more difficult with the added dimension of cultural and religious beliefs, values, and experiences of all those involved with or affected by the decision. Medical ethics has been defined as "moral principles that govern the practice of medicine" (Oxford Dictionaries 2016). Each individual's understanding of ethics is determined by his or her own personal values and beliefs and views of right and wrong. With the increased use of technology and health information for supporting the delivery of healthcare services, healthcare informatics (HI) has changed the landscape of healthcare and has introduced new situations resulting in the potential for new ethical situations to arise. More patient information is now available and used. This results in increased concerns for the rights of the patient for privacy, autonomy, and to be an active participant in decision-making regarding their care and in the management of their information (HIMSS 2016). There are many laws, regulations, and standards for protecting patient information related to privacy, security, unauthorized disclosure, use, and maintenance from creation to planned destruction. **Privacy** is patients' right to control the disclosure of their protected health information (AHIMA 2017). **Security** is the means to control access and protect the information from disclosure, accidental or intentional, to unauthorized individuals, and to protect from alteration, destruction, and loss (AHIMA 2017). Ethical situations and dilemmas surface when there is a conflict between one of the formal means of protection and what is perceived as the right thing to do in a given situation. An example of this is deciding how much and what information to give terminally ill patients. The patients have the right to their information and to be engaged in the decision-making process regarding their healthcare. Those close to the patients and even some of the providers of care might think it is more humane to withhold such information.

The pervasive use of informatics technologies and activities, such as electronic health records (EHRs), personal health records (PHRs), health information exchanges (HIEs), and data analytics, creates a complex ethical landscape that must be navigated carefully to avoid misuse, loss of protected health information (PHI), or unintentional destruction of PHI. HI can result in increased accessibility to patient information, but it is imperative that the integrity and quality of the data in the patients' health records be maintained. To navigate this shifting landscape, healthcare professionals must balance the benefits provided by the functionality of the informatics systems that facilitate sharing of patients' healthcare data from internal and external sources on multiple platforms in multiple locations with the need to protect patients' privacy and respect their autonomy (Layman 2008).

Defining Ethics

The word "ethics" is derived from the Greek word "ethos," which refers to letting your conscience guide you when deciding on the right or wrong action to take in a given situation (McWay 2021). **Ethics** has been defined as "a field of study that deals with moral principles, theories, and values; in healthcare, a formal decision-making process for dealing with the competing perspectives and obligations of the people who have an interest in a common problem" (AHIMA 2017). Humans constantly process data about the daily activities occurring in their surroundings into information used for making a multitude of ethical decisions. **Ethical decision-making** is a process where everyone must contemplate others' perspectives, even if they disagree on the issue (AHIMA 2017).

page 383

Ethical decisions have varying levels of difficulty. In general, the decision-making process must be internalized and ethical principles debated to reach equitable conclusions about the right and wrong actions for a given situation. Sometimes the decisions are relatively easy because the correct behaviors are dictated by outside forces. An example of easy ethical decision-making would be that anyone driving an automobile must obey the rules of the road or face the undesirable consequences of getting a traffic ticket. Most drivers realize that driving ethics is an example of common courtesy that has been translated into a set of laws designed to make the highways safer for everyone. These laws represent the acceptable choices, but it can be tempting to break the speed limit or run a red light to get to your destination faster. In healthcare, there are many situations that present ethical challenges and dilemmas for providers, such as beginning-of-life care, end-of-life care, and disagreements between patients and providers about treatment decisions (Ellison and Clack 2021). An example of an ethical decision related to the field of informatics would be the decision-making process for whether to include outliers in a data analysis project. There are no laws or specific requirements for including or excluding the outliers, and their presence can be attributed to several different things, such as errors in data collection or the presence of an extreme figure. These things, along with determining the influence of the outlier, require further analysis to determine whether the outlier is to be left in or removed (Murphy and Lau 2008). It would be simple to either keep it or remove it, but then the integrity of the statistical analysis might be in question and the analysis would be less useful. Thus, most ethical answers are not nearly as apparent in all situations and the decision-making can be a difficult process.

Values and Morals

The ethical choices made by individuals are based on their morals and values. **Morals** are the principles of appropriate conduct that an individual adopts (McWay 2021. Morals are grounded in the teachings of family, friends, religion, and other influential persons or groups such as significant teachers or workplace mentors. Morals can be conceptualized as the principles that people follow when deciding on what ethical actions they will take. **Values** are what ethics and morals are based on, the beliefs or standards on which judgment is based (McWay 2021). For example, if a childhood background included a hard-working parent, kind teacher, honest and loyal friends, or a strong religious framework, then these will likely be the values drawn upon when one makes moral decisions as an adult.

page 384

An individual's moral compass, one's capacity to assess right from wrong and perform accordingly, is strongly influenced by the groups to which they belong, the ethics of their workplace, and where they live. Individuals tend to norm to the groups' standards by bending their personal moral philosophy toward the group's moral philosophy to be accepted, which could have undesirable effects in the long run. Examples include a co-worker clocking in for another co-worker when they are late for work, or falsifying a favorite doctor's delinquent record list to gain acceptance.

Several widely used ethical concepts especially relevant to HI are autonomy, beneficence, nonmaleficence, and justice, which are discussed in this chapter (Budinger and Budinger 2006).

Autonomy

Autonomy refers to the freedom of individuals to choose their own actions. Autonomy is "a core ethical principle centered on the individual's right to self-determination that includes respect for the individual; in clinical applications, the patient's right to determine what does or does not happen to him or her in terms of healthcare" (AHIMA 2017). Autonomy would be exercised when obtaining informed consent from a patient for surgery. For the conditions of autonomy to be met, the preoperative patient would be given enough information about the surgery and its consequences, be able to understand that information, and cognitively reason to decide to sign the consent form for the surgery (Layman 2008). A related concept is **confidentiality,** protection against unauthorized access or unintended disclosure of the patient's written or verbal information (45 CFR 164.304 2013), which health information and HI professionals must protect. This could be done by restricting access to the PHI to include only those who have the proper authorization or right to access the information. Technically, confidentiality could involve the use of passwords and biometric user authentication for log-ins and having updated policies and procedures that are consistent with Health Insurance Portability and Accountability Act of 1996 (HIPAA) Privacy and Security regulations.

Beneficence

Beneficence is doing good for others. Beneficence, an act of goodness or kindness, implies that HI professionals have a duty to assist patients toward positive outcomes (Layman 2008). This principle is embodied in the Hippocratic Oath that physicians take on graduation, which requires them to treat patients in a manner that maximizes the patients' positive health outcomes (McWay 2021). In the field of HI, this could include verifying that the patient's information is being accessed only by the patient and others who will use this information for treatment, payment, or healthcare operations purposes that benefit the patient and those who have been authorized access. Beneficence is also being practiced when using available data to support clinical decision-making and the provisioning of quality healthcare (Cellucci et al. 2011). The increased use of self-monitoring devices for diabetics, telehealth psychiatric consultations for residents of remote areas, and the ability to print out discharge instructions are all beneficial to the patient (Layman 2008).

Nonmaleficence

Nonmaleficence is doing no harm to others. Nonmaleficence, protecting from wrongdoing, requires preventing access to the patient's data that could harm the patient (AHIMA 2017). An example is testing to ensure that electronic data are transmitted accurately between the laboratory computer and the **electronic health record (EHR)**. Nonmaleficence would include preventing harm to patients by retesting clinical decision support algorithms after system upgrades or security patches. Additionally, backup procedures must be in place to ensure business continuity in the event the system goes down, and antivirus programs must be kept up to date to prevent potentially harmful destruction of clinical data and ensure that recovery can be effected when needed.

Justice

Justice is an ethical principle that includes fairness and equality, and is displayed by treating everyone fairly and equally at all times (Olden 2019). Equal treatment can be demonstrated by maintaining everyone's PHI in a confidential and secure environment that is HIPAA compliant (International Medical Informatics Association 2016). Another

example of treating patients in a fair and equal manner is providing digital access via smartphones to facilitate a low-income patient's access to the internet and, consequently, portals and information. The application of justice is important in the ethical decision-making process, which is discussed in greater detail later in this chapter.

Legal Influence

At a minimum, members of society must abide by the legal rules set forth by that society. Thus, the laws of the land influence our ethical choices. In particular, the laws are the foundation for the development of the principles surrounding the ethical use of health information and technology in healthcare. When a circumstance is found to be in violation of the law or other regulation, it is not likely to be an ethical situation. The laws briefly mentioned here are covered in more detail in other chapters.

HIPAA

The Health Insurance Portability and Accountability Act of 1996 (HIPAA) is a package of federal legislation that, in part, assures privacy and security of health information (AHIMA 2017). HIPAA is complex and comprehensive, covering many different facets of both patient health information and the use of technology for creating, maintaining, accessing, disclosing, storing, using, and transmitting it. Consideration must also be given to the type of technology used with increased use of such things as mobile devices and internet computing.

HITECH

The Health Information and Technology for Economic and Clinical Health Act (HITECH), a part of the American Recovery and Reinvestment Act of 2009 (ARRA), is legislation designed to support the adoption of technology that will result in the meaningful use of patient health information (AHIMA 2017). Privacy, security, and the electronic exchange of data are key components in the act to reach the meaningful use goal. The electronic exchange of healthcare data involves electronic transfer among various entities.

GINA

Title II of the Genetic Information Nondiscrimination Act of 2008 (GINA), a federal law that took effect in 2009, was enacted to prohibit discrimination based on genetic information for health coverage and employment (HHS 2009). For GINA, due to the nature of the information obtained in genetic testing, genetic information was defined as information about an individual's genetic tests, their family's genetic tests, or their family medical history. Different from most other medical tests, identifying a genetic marker in one person also identifies it for some family members as well.

HIPAA Omnibus Rule

The HIPAA Omnibus Rule was enacted in 2013 to strengthen privacy and security protections that were established in the initial HIPAA (HHS 2009). The Omnibus Rule primarily addresses the privacy protections, new rights for individuals related to their health information, and increased government ability for enforcement. All of the changes are consistent with the other laws, HITECH and GINA, that have been enacted since HIPAA first went into effect and with the increasing digital world.

Even when there are laws or rules in place to govern a situation, ethical issues might still arise when the law is not explicit in a given circumstance. The nature of health information and the increased use of health information technology can foster the development of ethical situations where existing laws might be in conflict with personal values. Examples of these instances include where some are in favor of sharing PHI when a communicable disease is involved (for example, contact tracing for COVID-19 cases)

or sharing genetic information to alert other family members of their risk for potential conditions. With the increased use of technology throughout the healthcare enterprise, there are many more employees who have access to health information via electronic systems, but they might not have the right or need to know the PHI.

Ethical Standards of Practices

Ethical dilemmas occur when there are differing perspectives among the individuals involved in the situation including the patient, patient's family, and friends to the countless healthcare professionals who treat and provide services from all facets of supporting the delivery of healthcare. A **code of ethics** is "a statement of ethical principles regarding business practices and professional behavior" (AHIMA 2017). Professional associations and societies support a sound ethical framework that members are obligated to follow. Many of the healthcare codes of ethics include a statement referring to protecting patient privacy and respecting the patient. The AHIMA Code of Ethics is designed to reflect the obligation of the health information management (HIM) professional to demonstrate actions that reflect values and ethical behavior (AHIMA 2019). Figure 13.1 provides the principles of the AHIMA Code of Ethics. Information regarding the interpretation of specific behaviors expected by HIM professionals can be found in the full AHIMA Code of Ethics available on the AHIMA website.

Figure 13.1. Principles of the AHIMA Code of Ethics

The following principles are based on the core values of the American Health Information Management Association and apply to all AHIMA members, non-members, Commission on Certification for Health Informatics and Health Information Management (CCHIIM) certifications, and students.

1. Advocate, uphold, and defend the consumer's right to privacy and the doctrine of confidentiality in the use and disclosure of information.
2. Put service and the health and welfare of persons before self-interest and conduct oneself in the practice of the profession so as to bring honor to oneself, their peers, and to the health information management profession.
3. Preserve, protect, and secure personal health information in any form or medium and hold in the highest regard health information and other information of a confidential nature obtained in an official capacity, taking into account the applicable statutes and regulations.
4. Refuse to participate in or conceal unethical practices or procedures and report such practices.
5. Use technology, data, and information resources in the way they are intended to be used.
6. Advocate for appropriate uses of information resources across the healthcare ecosystem.
7. Recruit and mentor students, peers, and colleagues to develop and strengthen professional workforce.
8. Represent the profession to the public in a positive manner.
9. Advance health information management knowledge and practice through continuing education, research, publications, and presentations.
10. Perform honorably health information management association responsibilities, either appointed or elected, and preserve the confidentiality of any privileged information made known in any official capacity.
11. State truthfully and accurately one's credentials, professional education, and experiences.
12. Facilitate interdisciplinary collaboration in situations supporting ethical health information principles.
13. Respect the inherent dignity and worth of every person.

AHIMA also sets professional ethical standards for coding professionals through the AHIMA Standards of Ethical Coding (AHIMA 2016). Figure 13.2 provides the principles of the AHIMA Standards of Ethical Coding. AHIMA also has ethical standards for other professions, such as the Ethical Standards for Clinical Documentation Integrity Professionals, which can also be found on the AHIMA website (AHIMA 2020).

Another professional code of ethics directly associated with HI is that of the American Medical Informatics Association (AMIA), whose members are engaged in transforming healthcare through science, educational opportunities, and biomedical and health informatics practice (AMIA 2016). Figure 13.3 shows AMIA's Code of Professional and Ethical Conduct (Petersen et al. 2018).

Figure 13.2. Principles of the AHIMA Standards of Ethical Coding

1. Apply accurate, complete, and consistent coding practices that yield quality data.
2. Gather and report all data required for internal and external reporting, in accordance with applicable requirements and data set definitions.
3. Assign and report, in any format, only the codes and data that are clearly and consistently supported by health record documentation in accordance with applicable code set and abstraction
4. conventions and requirements.
5. Query and/or consult as needed with the provider for clarification and additional documentation prior to final code assignment in accordance with acceptable healthcare industry practices.
6. Refuse to participate in, support, or change reported data and/or narrative titles, billing data, clinical documentation practices, or any coding related activities intended to skew or misrepresent data and their meaning that do not comply with requirements.
7. Facilitate, advocate, and collaborate with healthcare professionals in the pursuit of accurate, complete, and reliable coded data and in situations that support ethical coding practices.
8. Advance coding knowledge and practice through continuing education, including but not limited to meeting continuing education requirements.
9. Maintain the confidentiality of protected health information in accordance with the Code of Ethics.
10. Refuse to participate in the development of coding and coding related technology that is not designed in accordance with requirements.
11. Demonstrate behavior that reflects integrity, shows a commitment to ethical and legal coding practices, and fosters trust in professional activities.
12. Refuse to participate in and/or conceal unethical coding, data abstraction, query practices, or any inappropriate activities related to coding and address any perceived unethical coding related practices.

Figure 13.3. Principles of Professional and Ethical Conduct for AMIA members

I. Key ethical guidelines regarding patients, guardians, and their authorized representatives (called here collectively "patients")

AMIA members involved in patient care should:

A. Recognize that patients and their loved ones and caregivers have the right to know about the existence and use of electronic records containing their personal healthcare information, and have the right to create and maintain their own personal health records and manage personal health information using a variety of platforms including mobile devices. In this context AMIA members should:

1. Not mislead patients about the collection, use, or communication of their healthcare information.
2. Enable and — as appropriate, within reason and the scope of their position and in accord with independent ethical and legal standards — facilitate patients' rights and ability to access, review, and correct their electronic health information.
3. Recognize that patient-provided/generated health data, such as those collected on mobile devices, deserve the same diligence and protection as biomedical and health data gathered in the process of providing health care.

B. Advocate and work as appropriate to ensure that protected health information (PHI), personally identifiable information (PII), and other biomedical data are transmitted, acquired, recorded, stored, maintained, analyzed, and communicated in an appropriately safe, reliable, secure, and confidential manner, and that such data management is consistent with applicable laws, local privacy and security policies, and accepted informatics standards.

C. Never knowingly disclose PHI, PII, or biomedical or health data in violation of legal requirements or accepted local confidentiality practices, or in ways that are inconsistent with the explanation of data disclosure and use to the patient. AMIA members should understand that inappropriate disclosure of biomedical information can cause harm, and so should work to prevent such disclosures. AMIA members should avoid acquiring data through means that run the risk of, or fail to prevent, inappropriate disclosure. Likewise, even if an action does not involve disclosure, one should not use — or through negligence permit the use of — patient information and data in ways inconsistent with the stated purposes, goals, or intentions of the patient or organization responsible for these data, except as appropriate for public health, previously approved and communicated research uses, or reporting as required under the law.

II. Key ethical guidelines regarding colleagues

AMIA members should:

A. Endeavor, as appropriate, to support and foster colleagues' and/or team members' work, in a timely, respectful, and conscientious way to support their roles in healthcare and/or research and education.

B. Support and foster the efforts of patients to be actively involved in the collection, management, and curation of their health data.

C. Advise colleagues and others, as appropriate, about actual or potential information or systems issues (including system flaws, bugs, usability issues, etc.) that negatively affect patient safety, privacy, data security, or outcomes or could hinder colleagues' ability to delegate responsibilities to patients, other colleagues, involved institutions, or other stakeholders.

D. If a leader, an AMIA member should:

1. Be familiar with these guidelines and their applicability to their practice, unit, or organization.
2. Communicate as appropriate about these ethical guidelines to those they lead.
3. Strive to promote familiarity with, and use of, these ethical guidelines.

III. Key ethical guidelines regarding institutions, employers, business partners, and clients (called here collectively "employers")

AMIA members should:

A. Understand their duties and obligations to current and former employers and fulfill them to the best of their abilities within the bounds of ethical and legal norms.

B. Understand and appreciate that employers have legal and ethical rights and obligations, including those related to intellectual property. Understand and respect the obligations of their employers, and comply with local policies and procedures to the extent that they do not violate ethical and legal norms. Consider the tradeoffs that occur with the configuration and use of technologies (e.g., decision support systems) before implementation, and monitor and manage results when the optimal approach is unclear.

C. Inform the employer and act in accordance with ethico-legal mandates and patient rights when employer actions, policies, or procedures would violate ethical or legal obligations, contracts, or other agreements made with patients. Maintain a safe and high-quality environment even while implementing innovation, recognizing that all changes in a complex adaptive environment generate unanticipated consequences and potential harm.

IV. Key ethical guidelines regarding society and regarding research

AMIA members involved in research should:

A. Be aware of the Declaration of Helsinki (Ethical Principles for Medical Research Involving Human Subjects), which should guide all human subject research, including research that involves users of informatics tools and interventions as human subjects (e.g., workflow analysis studies, clinical decision support systems analysis, patient care innovations, analysis, etc.). Recognize that duty and care to colleagues exist regardless of whether such responsibilities are acknowledged by institutional review boards, vendors, and others involved in informatics activities.

B. Be mindful and respectful of the social or public health implications of their work, ensuring that the greatest good for society is balanced by ethical obligations to individual patients.

C. Avoid any plagiarism or self-plagiarism or other misrepresentations of the truth in the publication of research and other work.

D. Disseminate new knowledge — both positive and negative — expeditiously, to allow the field to advance and to permit others to take advantage of novel discoveries to improve patient care.
E. Strive as appropriate in the context of one's position to foster the generation of knowledge and biomedical advances through appropriate support for ethical and institutionally approved research efforts facilitated through informed consent and disclosure processes and procedures, particularly when third-party entities not meeting the definition of business associates are involved.
F. Know and abide by the applicable governmental regulations and local policies that define ethical research in their professional environment.

V. General professional and ethical guidelines

AMIA members should:

A. Maintain competence as informatics professionals:
1. Obtain applicable continuing education and be dedicated to a culture of lifelong learning and improvement;
2. Recognize technical and ethical limitations and seek consultation when needed, particularly in ethically conflicting situations;
3. Contribute to the education and mentoring of students, junior members, and others, as appropriate;
4. Promote a culture of inclusivity in their work and professional conduct.

B. Strive to encourage the adoption of informatics approaches supported by adequate evidence to improve health and healthcare; and to encourage and support efforts to improve the amount and quality of such evidence.
C. Be mindful that their work and actions reflect on the profession and on AMIA.

Source: Petersen et al. (2018)

As can be seen with these statements, patient and other confidential information in any form must be protected, and the concern for the individual is of utmost importance.

Policies and Procedures for Ethics Related to Informatics

To operationalize the laws and other regulations, consistent codes of ethics such as the AHIMA Code of Ethics, AMIA's Code of Professional and Ethical Conduct, or another code of ethics, policies and procedures must be developed that serve to protect the privacy, security, and confidentiality of PHI, and act as a deterrent to unethical workplace practices. These policies and procedures must facilitate personal freedom as it relates to technology use (autonomy), ensure technology is used for good (beneficence), and prevent technology from causing harm to others (nonmaleficence; Anderson 2004). This requires balancing the need to give access against the need to secure the information systems. The following example shows the number and types of informatics policies and procedures that are used to create an ethical healthcare workplace by explaining the process of getting a new hire's technical work environment set up, which illustrates the minimum necessary policies and procedures to ethically use HI.

A new employee, John, has been hired for the HIM department. To begin, after completing the new hire paperwork and passing the required background check, he is assigned a job title and a job role, and a department in the Human Resources database

system. When John arrives for his first day at work, he is unable to enter the building until he goes to the security office, presents identification, and has a plastic identification badge issued. Security personnel call his department, and someone comes to escort him to the HIM department where he is required to test his badge by swiping it on the electronic badge reader for the HIM department's door.

After being shown to his cubicle, John finds that he does not have a computer system on his desk. At this point, the HIM department manager follows the procedure for ordering a new computer for the new employee with the company's approved computer configuration. The manager knows that this standardization is needed to more easily maintain the facility's systems (Abdelhak et al. 2011). The IT department places the order, specifying an up-to-date system. When the new employee's computer arrives at the facility, the type and characteristics of the machine are recorded to be used for later maintenance and inventory tracking. Next, the new computer is assigned an internal asset code, which is etched into a plastic plaque and attached to the machine as part of the security procedure to uniquely identify the machine, to prevent theft of the machine, and to allow asset tracking. The facility will use a "standard build" that is a set of specific software (for example, Microsoft Word, Outlook email, and a web browser), which will be installed on all new machines, and specific tests will be run to ensure that the machine is working according to specifications (Denic´ et al. 2014).

When the machine is ready, it is taken to the department and placed on John's desk. Unfortunately, he finds that he cannot log on to the machine until he goes to his boss and gets his log-in and temporary password, which is a standard security procedure. Using his network ID and password, John logs onto the network and is then able to access his Windows system, but his email and web browser are giving errors related to not having an internet connection. Due to the number of healthcare breaches occurring over the internet, his boss must complete a second form stating that he needs the internet access to perform his work activities (HHS 2014; Butler 2015). After John receives internet access and has a network connection physically set up on his computer, hiss manager tells him to go ahead and complete the mandatory online HIPAA training on privacy and security, and to take the other mandatory classes on facility security, ethics, and the use of emergency codes. He is also given the HIM department policy and procedure book to read, and he must sign a form stating that he has read it. John learns that use of employer-owned property for personal use is prohibited. He will email his friends after work.

While completing this training, John has some technical issue with his email and calls support, who follows security procedures by directing him to read the asset tracking code from the machine and to tell them his network ID. While using his computer, John notices that some websites he tries to access are blocked because they do not have valid security certificates, and he observes that virus tracking software and routine backups are running on the machine (Abdelhak et al. 2011). John has completed all these preliminary steps and is ready to get started on his work assignments. His manager tells him that he will be using three software packages at work, which are an electronic health record, a chart tracking system, and a system that tracks record deficiencies. Each of these systems requires further training and each system has its own procedure for security access and for logging on with different usernames and passwords. John learns that his access levels for these systems will be determined by his assigned role at work, which is called role-based access control, and that this security feature helps IT manage and track system usage.

In the lunchroom, John ends up sitting at a table with Jane, the IT employee who set up his new machine. Jane tells him that he should feel confident that IT is using technology for the benefit of all employees and patients. Specifically, the IT department stores the servers in a restricted area that requires biometric access, they perform regular backups of

servers, and their maintenance schedule includes regular updates to all system hardware and software, as well as an annual disaster recovery drill. They also have offsite storage of their backups, and uninterruptable power systems to provide power if the system goes down. In addition, they monitor the audit logs on all servers to determine who has accessed the systems and what they have done.

Jane notices that John has brought a personal tablet to work and is using a smartphone to read his work email at meetings. She tells him that there is a policy requiring him to report any personal device usage at work to his boss and that he will have to come into the IT department to have his devices physically checked or he will not be able to continue bringing the devices to work (Thomas 2007; HIMSS 2011; Tan and Aguilar 2013; Imgraben et al. 2014). She also mentions that he needs to be encrypting all his work email to protect the PHI. At this point, Jane has to leave to return to a workshop on implementing secure patient-centered PHI, which is part of her department's mandatory professional development training designed to keep all employees up to date on the latest technology. As John walks back to his desk, he is tired but highly impressed by the numerous policies and procedures in place to protect the employees and patients' data that form the ethical framework of this workplace.

This example shows the number and types of policies and procedures required for PHI for a facility to be compliant with the privacy and security laws. Professionals working in this environment are personally obligated to conduct themselves with ethical behavior, and managers are accountable for assuring that the employees have the tools in the form of the policies and procedures and a culture that respects ethical behavior to promote ethical decision-making.

The Ethics Committee

Traditional functions of a healthcare ethics committee are to provide consultation when ethical issues arise, generate and revise policies and procedures for clinical ethics, and facilitate entity educational activities for employees (Pearlman 2013). The committee membership may vary based on the emphasis and function of the committee; that is, a clinical ethics committee will likely be comprised of mainly physicians, whereas a more comprehensive hospital ethics committee will have a more diverse membership from across the facility. More facilities are expanding the functions of their ethics committees to cover issues related to both clinical and organizational ethics situations. In addition to the patient and provision of care-related goals of the traditional clinical ethics committee, the current trend is expanding the ethics committee goals to include:

- "Integrating ethics throughout the health care institution from the bedside to the boardroom,
- Ensuring that systems and processes contribute to/do not interfere with ethical practices, and
- Promoting ethical leadership behaviors, such as explaining the values that underlie decisions, stressing the importance of ethics, and promoting transparency in decision making" (Pearlman 2013).

The inclusion of goals addressing ethical situations across the organization's personnel, their systems, and their processes is significant. These items also suggest that the leadership in an organization must strive for an institutional culture that values ethical decision-making and demonstrates ethical behavior.

Check Your Understanding 13.1

Answer the following questions.

1. Which field of study deals with principles, theories, and values related to our moral choices?
   a. Ethics
   b. Values
   c. Nonmaleficence
   d. Justice
2. Which term refers to the freedom of individuals to choose their own actions?
   a. Autonomy
   b. Beneficence
   c. Nonmaleficence
   d. Justice
3. Treating all people fairly is referred to as
   a. Autonomy
   b. Beneficence
   c. Nonmaleficence
   d. Justice
4. Which of the following is an example of beneficence?
   a. Verifying patient's identity during admission
   b. Letting your friend use your password
   c. Locking your computer when you go to lunch
   d. Getting an informed-consent form signed
5. Which of the following terms means doing no harm to others?
   a. Beneficence
   b. Autonomy
   c. Nonmaleficence
   d. Justice

Potential Ethical Issues in Healthcare

Numerous ethical dilemmas arise during the provisioning of healthcare. For example, how can healthcare professionals use health information ethically? How can they protect the privacy and security of the patients' data while providing access to patients and providers via EHRs, PHRs, and mobile devices? Additional ethical concerns surround data sharing through the internal processes of release of information, the secondary release of information, and the external collaborations inherent in the health data exchange systems. **Secondary release of information,** or redisclosure, is releasing the health information that was originally obtained per a request from an external facility (AHIMA 2013). A growing body of research on health literacy raises concerns about what patients really take away from their perusals of public websites, or from the handouts they are given at their doctor's office. Finally, the use of social networks is pervasive, and this creates a gray area for employers because information from work shared via social networks can do irreparable harm in minutes, and it is hard to retract all copies.

Provisioning of Care

Facilities have a responsibility to establish a corporate culture that rewards ethical behavior and applies sanctions for unethical activities (Lefkowitz 2006). When the provision of healthcare meant that the doctor and the patient were in the same room, there was a strong connection between the healthcare provider and the patient. Similarly, the

healthcare providers' daily ethical decision-making was influenced by the everyday interactions with co-workers, managers, and the top leadership of the facility. Although technology has provided many beneficial advances in healthcare using robotic surgery, telehealth, patient portals, and remote management of patients' chronic diseases via smart devices, unfortunately, it has also created the potential for new and different kinds of ethical dilemmas. Indeed, many healthcare workers will never set foot inside the facility, but they will manage the patients' data (through cloud computing or remote data storage centers) or their care from a distance (for example, telehealth).

In the past, workers commuted to work, entered the healthcare facility, and were exposed to the corporate ethical culture of their company's top management. They interacted face-to-face with patients and staff, attended department meetings in person, and had a chance to observe firsthand the demeanor of their department heads as they conducted business. Consequently, a facility with strong ethical leaders and a culture that fostered autonomy, beneficence, equality, charity, and justice provided a strong moral compass for their employees to follow. In this era of globalization and virtual private networks, there are more and more health employees telecommuting from home, from other states, or even from other countries, as is common in the field of teleradiology. One researcher argues that the old concept of employees wanting to be seen as good corporate citizens has gone by the wayside as virtual healthcare workers and electronic medical record systems become the norm (Lefkowitz 2006). Consequently, although the ethical behavior practiced by top management still strongly influences their employees' ethical behavior, the effectiveness of this positive influence is diffused when the face-to-face element between employer and employee is lacking.

Ethical Use of Health Information

Using health information ethically means honoring the ethical concepts of autonomy, beneficence, nonmaleficence, and justice. In the workplace, HI professionals must deal with complex ethical issues in accordance with the standards of practice clearly outlined in the AHIMA Code of Ethics. Unfortunately, everyday workplace ethical dilemmas are multifaceted and usually do not have clear-cut answers. Although it is ethically and morally correct to report upcoding for the purpose of higher reimbursement, the reality is that the **whistleblower**, one who reports a wrongdoing, might be asked to seek other employment, which is a misuse of justice (Crawford 2011). Similarly, failing to report delinquent records to stay below the required Joint Commission delinquency rate is dishonest, but reporting the oversight to your manager might cause you harm if the manager is the one who is generating the delinquency reports. There are also situations where the person behaving unethically is not aware of the rules governing the situation, which is why healthcare facilities require employees to remain up to date by completing annual professional development classes on relevant rules and regulations for their jobs. Indeed, the principles of ethics and risk management are interrelated because the facility is responsible for maintaining current policies and procedures and documenting that employees have received training on those policies and procedures.

The use of EHRs has introduced functionality for entering information into a patient's record where the information might not be as precise or as specific to an individual patient as intended while still being technically correct. One example of this is the copy/paste functionality, which includes copying material from one document such as a record from a previous hospital stay and pasting it into a patient's EHR or copying from one part of a record to another without deleting from the original location (AHIMA 2017). This is a shortcut for documenting but can have consequences primarily related to the integrity of the record. When blocks of information are copied and pasted, there could be parts of it that are inaccurate for the current encounter and the possibility of

contradictions with other information in the record could occur easily. There are also some situations that could result in legal issues, such as the pasted information having a higher reimbursement than what the patient actually was being treated for; this would then be considered fraud. An example of this is a patient with two hospitalizations just months apart, both for respiratory conditions. On the first visit the patient was treated for a bacterial pneumonia and on the second visit was treated for bronchitis. Both conditions had many of the same symptoms, tests, and treatments, but, in the end, they were different conditions. Copying and pasting from the first visit for the second visit, without deleting symptoms, diagnostics, and treatments that were only done in the pneumonia visit could inadvertently trigger the reimbursement rate for pneumonia rather than bronchitis. Knowingly providing documentation that is not current or specific to a given situation is not in the best interest of the healthcare of the patient. Even when fraud is not an issue, once it is recognized that the copy and paste has been done, the credibility of the record will be in question.

Autonomy and confidentiality of patient information can be hard to preserve in this digital era where smartphones are used to transmit photographs of patients to their doctors, and hospitals have surveillance cameras in the halls (van der Rijt and Hoffman 2015). Similarly, administering beneficence in the context of making choices for health information technologies (HITs) that benefit patients has become a balancing act. There are evidence-based, long-term benefits from funding multiyear high-dollar electronic record system implementations, such as a reduction in medical record errors, more complete and more specific documentation for billing, and a facilitated release of information cycle. However, a study by the Joint Commission on sentinel events associated with the implementation of computerized provider order entry (CPOE) reflect that HIT has been a contributing factor in medical errors. A **sentinel event** is "an unexpected occurrence involving death or serious physical or psychological injury or risk" (AHIMA 2017). Findings of the Joint Commission review of 3,375 sentinel events occurring between January 1, 2010, and June 30, 2013, indicate the following:

- A total of 120 sentinel events identified HIT as a contributing factor.
- Fifteen different types of events were associated with the most common events: (1) medication errors, (2) wrong-site surgery (wrong side/site, wrong procedure performed or on wrong patient), and (3) delays in treatment.
- Multiple contributing factors to any one sentinel event with 305 HIT–related factors identified across the 120 HIT–related sentinel events with the following being the most common:
- Human-computer interface, 33 percent
- Workflow and communication, 24 percent
- Clinical content, 23 percent
- There were 120 sentinel events; however, 147 medical devices were involved because more than one HIT device could have been involved in a sentinel event.
- Sixty-six percent of the sentinel events involved EHRs or some component of the EHR (includes CPOE, pharmacy, and clinical documentation systems [for example, progress notes], eMARs, and clinical decision support; HHS 2015).

The Joint Commission study illustrates that there are associated risks when incorporating HIT, although the majority of these cases caused only minor health issues or delayed appropriate treatment. Another study that reported much more dire results attributed to CPOE system implementation, completed in 2005, brought attention to the issue. The study reported:

> Of 1,942 children admitted for specialized care during the study period, 75 died, a 3.86% mortality rate. Further analysis revealed the mortality rate to be 2.80% (39 of 1,394) before CPOE and 6.57% (36 of 548) after CPOE was implemented. This finding was an unexpected increase in mortality coincident with CPOE implementation. The study recognized that CPOE technology has the capability to reduce human error but the unexpected finding indicates continued evaluation of the mortality effects and medication error rates especially for children dependent on time-sensitive therapies. (Han et al. 2005).

The Han et al. report brings attention to the serious nature of outcomes, including mortality, that can occur when there are errors made in using the CPOE systems, which is in conflict with what has been touted as a benefit to using CPOE. Moreover, malpractice claims related to EHR risks have resulted from system issues such as poor system design (10percent), system or technology failure (9 percent), alert or alarm issues (seven percent), electronic data routing failure (6 percent), lack of space for documentation (4 percent), or fragmented EHRs (3 percent). User factors have also caused risk issues, with the top three reported issues being incorrectly entered information (16 percent), problems with hybrid health records or EHR conversion (15 percent), and issues with preloading information (13 percent; Crysts 2016, 15).

All of these studies could be considered a violation of ethical principles when considering the rights of patients, expectation of sound medical, patient-centered outcomes, and ensuring effective systems and practices to support the delivery of care and patient privacy.

Protected Health Information

The ethical concept of autonomy, also called freedom of choice, is closely related to the legal concept of privacy. Privacy in healthcare is considered to be the right of patients to control disclosure of their PHI (AHIMA 2017). Historically, there was a limited concept of patient privacy, other than that related to the patients' medical records, and it was not unusual to see patients' names on the door of their room and also on a chalkboard in the hall by the nurses' station so that visitors and staff could more easily locate them. For many years, nurses' stations had a bulletin board that listed the patients' names, room numbers, and physicians, and the procedures they were having (for example, upper GI or gallbladder surgery) so that staff could keep track of who had gone to surgery or radiology and who still needed to be transported. When a patient was transported to x-ray, the paper chart was placed beside the patient on the stretcher, which meant that the chart was somewhat accessible to anyone walking by who wanted to take a look at that information.

In the digital era, most PHI is collected, transmitted, and stored electronically, so there is no one "door" that can be locked to control access to the data. For example, while the architecture of the internet facilitates the remote monitoring of a sudden infant death syndrome baby and the subsequent data transmission to cloud storage for later download to the provider's smartphone, it also makes it almost impossible to guarantee that the patient's private data are safe from unauthorized disclosure (Levine 2015).

Genetics

Compared to most PHI, additional ethical questions arise with genetic information due to the nature of the information that is obtained with the testing (NSW Health 2016). Genetic data are different because they engender the concept of shared ownership of genetic information, due to its familial nature. Moreover, genetic testing is not always definitive because it identifies present or potential conditions but cannot determine when the condition might present itself or the severity of the condition if it does present. This

uncertainty about future conditions complicates current ethical decision-making. How the information is used also causes ethical dilemmas, such as when the genetic information is used for employment screening, for obtaining insurance, and for making reproductive decisions. Genetics information is valuable, with the potential to significantly improve health outcomes using preventive measures and personalized treatment based on genetic test results. However, caution must be exercised with the information, as illustrated in the following case study.

Case Study: Genetic Predictive Testing

Huntington's disease (HD) is a neurological progressive degenerative disease with no cure. Onset usually occurs between 30 and 50 years of age with symptoms of declining movement, cognition, and general ability to function, with respiratory conditions commonly the cause of death. It is an inherited disease. A child of a person who has the disease has a 50 percent chance of inheriting the faulty gene. Individuals over the age of 18 who have an affected parent or relative can obtain genetic predictive testing. The results of the test will reliably reveal whether they will develop the disease at some point in their life.

Mr. C is a 25-year-old male. His grandfather died from HD 10 years ago, which means that his mother has a 50 percent chance of developing HD. She was genetically tested and does have the faulty gene, which indicated that she will certainly develop HD at some time, and now Mr. C. is at a 50 percent risk of developing HD.

Mr. C. is an air traffic controller and he does not want to have the test. He likes his job and feels that he can carry out his duties appropriately for years whether he carries the faulty gene or not. His employer is not aware of the family medical history of HD.

The ethical questions at this point are as follows:

- Should Mr. C have the predictive testing? Identify both the pros and cons he should consider in making a decision.
- Do employers have the right to know family medical history, especially where public safety could be an issue? At what point do they have the right to this information? While the individual is still healthy, later when symptoms to start to appear, or never?
- Whose information is this? The mother's, Mr. C's, or both? Who can make the decision to release it?
- What if in this situation Mr. C. wanted the testing but his mother refused? His results might reveal the genetic status of the mother (if Mr. C has it, he must have inherited it from his mother).
- Consider the implications for Mr. C's reproductive choices. Does he tell the potential mother of his children and at what point in the relationship? If a child is conceived, should prenatal testing be done and what choices must be made (NSW Health 2016)?

This case study illustrates an ethical dilemma with a number of implications and decisions to be made by several individuals. There are no right or wrong actions, with much left to personal choice. For this reason, genetic

counseling is a common part of the genetic testing process. This case also illustrates the importance of protecting the PHI.

page 381

Electronic Health Records

EHRs are "an electronic record of health-related information on an individual that is created and managed by clinicians and staff across two or more healthcare organizations" (AHIMA 2017). The use of EHRs has made PHI documentation by physicians more legible and accessible, which can benefit patients by contributing to the reduction in medical errors, providing data integrity, and facilitating communication among interdisciplinary healthcare teams (Cheshire 2014). EHRs enable e-prescribing and computerized order entry, they provide alerts for physicians, and they are accessible from multiple locations (Cheshire 2014). Conversely, the confidentiality of the patient's e-PHI is harder to protect due to the increased accessibility, and many feel that the patient–provider relationships may suffer (Clark 2014; Cochran et al. 2015).

Today PHI is stored electronically and the virtual storage media makes it easier to steal. First, e-PHI can be accessed remotely using mobile devices, web-based PHRs, and patient portals, which greatly reduces the danger of being observed during the theft. Second, the thief is pilfering data from a computer, thus the sense of connectedness to stealing an actual person's healthcare data is diminished.

Healthcare entities that do attempt to prosecute malicious users soon realize that these unauthorized users believe they are not subject to the laws of society (Cross 2006). Researchers have coined the term **hacker ethics** to refer to the hackers' beliefs that restrictions on data are objectionable and that technology is a tool that should be used to perform hands-on actions that make nonhackers' private data publicly accessible while protecting the hacker's personal privacy (Kirkpatrick 2002). The size of this problem cannot be underestimated; hackers are not just a local issue involving one disgruntled IT employee in a garage. Hackers form global groups to share system vulnerabilities and system access codes on their websites. HIM and HI personnel should conduct a security audit to assess vulnerabilities. As a final note, there are some hackers called white hat hackers who use their abilities to help companies detect their vulnerabilities (Vaas 2007). Essentially, a white hat hacker breaks or hacks into protected systems and networks to assess the security and identify vulnerabilities.

To reduce security vulnerabilities, measures like antivirus software, role-based access control, passwords, data encryption, audit trails, and server security are needed to protect PHI against hackers (HIMSS 2011). The facility should develop a security profile to be used when selecting vendors. New computers should have a standard security configuration. Computer system inventories must be kept up to date so that the location and security measures in use on each system can be tracked; this will also assist in the timely replacement of older systems. Policies and procedures must address the level of system access to give the patient's family and visitors through wireless connections (HIMSS 2011). An example of how the EHR system facilitated the unauthorized access to a patient's information follows:

> A nurse practitioner who has privileges at a multi-hospital health care system and who is part of the system's organized health care arrangement which allows access to all patient records across the system impermissibly accessed the medical records of her ex-husband. Following a determination by the OCR, the covered entity: terminated the nurse practitioner's access to its electronic records system; reported the nurse practitioner's conduct to the appropriate licensing authority; and provided the nurse practitioner with remedial Privacy Rule training. (HHS 2016)

page 382

This example illustrates the importance of strong policies and procedures, appropriate access levels granted, and the employees' education. In this particular situation, a review of the practice of granting access to all patient records across the systems should be conducted to determine if that is necessary, and policies and procedures to address whatever access is provided should be updated.

Personal Health Records

The Centers for Medicare and Medicaid Services (CMS) federal incentive Stage 2 Meaningful Use program promotes the meaningful use of EHRs and recommends patient engagement through access to their PHI through a patient portal or by obtaining a copy of their personal records from their provider (HHS, CMS 2012). A **personal health record (PHR)** is a paper or electronic record maintained by patients for the purpose of collecting and sharing their data (AHIMA 2017). There are two types of PHRs. A stand-alone PHR is generated by the patient from their own records with information stored on a personal computer or the internet. These have the capability to accept information from external sources. Patients have control of what to include and may include such things as diet or exercise logs. They also control with whom to share the information. A tethered or connected PHR is linked to a specific organization's EHR system. Patients have access via a web portal or a mobile app (HealthIT 2016). PHRs provide patients with more autonomy by providing greater access to their health data and greater ownership of those data. In PHR portals, patients can view health information about diseases, examine their diagnostic results, email providers, or schedule appointments (Tang et al. 2006).

A PHR can be hosted on a vendor's website where the patient logs in and enters the information (stand-alone) or it can be integrated into the hospitals' system such that part of the data is entered automatically by the hospital's systems and part of the data is entered by the patient (Househ et al. 2014). The patient portal functionality of an EHR is controlled by the healthcare organization. The patient has read-only access to view only the designated sections. Email and appointment scheduling might also be available, similar to the PHR portal.

There are two major barriers to increasing PHR usage. First, the absence of a standard design for PHR systems or a standard PHR data set for data transmission has led to a lack of interoperability among PHRs and other healthcare systems (Studeny and Coustasse 2014). Second, patients are concerned about the privacy of their PHR data.

The following are examples of PHR software systems:

- Epic's MyChart
- Department of Veteran Affairs' My HealtheVet
- Health Companion

One technical challenge with PHRs is ensuring the privacy and security of PHI in the PHR systems due to the lack of standard designs and data sets. Moreover, because so much computer code has to be generated to move PHR data between these disparate healthcare systems, and because there is no standard design or standard transmission data set, there are concerns about data accuracy. Physicians are especially cautious about the legality of making medical decisions based in part on data recorded by the patient in their PHR (Studeny and Coustasse 2014).

Ethical challenges include the concerns for maintaining complete and current information and determining what is the level of responsibility for maintaining the PHR. Although physicians have a concern for using patient-generated data, patients have a similar concern about the information populated for them in their PHR. Concerns include

whether the test results and medication records are up to date, if all their reports are there, and if there is a possibility of information for another patient ending up in their record.

Release of Information

Release of information (ROI) refers to the process of disclosing PHI from health records (AHIMA 2017). The ROI process requires corresponding policies and procedures designed to restrict access to the patient's information to those individuals who are authorized by the patient to receive the information to be in compliance with existing laws. An organization is bound to use systems and processes that assure that improper disclosure is prevented. The following is an example in which a health plan had a computer flaw that resulted in mailing EOBs (Explanation of Benefits) to the wrong person and the corrective action taken:

> A HMO sent an EOB by mail to a complainant's unauthorized family member. An investigation by the OCR identified a flaw in the health plan's computer system that put the PHI of approximately 2,000 families at risk of disclosure in violation of HIPAA Security Rule. The corrective action required the insurer to correct the flaw in its computer system, review all transactions for a six month period and to correct all corrupted patient information. (HHS 2017.)

With an organization's responsibility to use systems that assure only proper disclosures are made, it is imperative that system testing be performed periodically to make sure they are working as intended. Correcting corrupted information can be costly in terms of time, money, and reputation.

Health Information Exchange

To receive incentive money from the HITECH Act, a facility must demonstrate use of electronic records for data sharing. This can be accomplished with a **health information exchange (HIE)**, which expedites data sharing between healthcare systems located in different facilities (Vest and Gamm 2010). HIEs can contain massive amounts of data with the potential for meaningful biomedical and public health research. The concepts of beneficence, to do good, and nonmaleficence, to do no harm, are in conflict in this instance. The research would benefit the health of individuals and the populace, but to do no harm would require both that the data in the exchange be protected from unauthorized use and that consent be given from each individual in the exchange for the specific use of their data and for researchers who will require access to the systems. HIE systems at this time do not have the capacity to be research compatible, and adding this capacity would be very costly and legally challenging (Mercuri 2010).

Social Networking

Social networking refers to social communications conducted using internet-based **social media** tools such as email, text messaging, Facebook, Twitter, Instagram LinkedIn, and YouTube. Social media use is becoming pervasive in our society. As of January 2022, there were 436 million Twitter users and 2.9 billion active Facebook users (Statista 2022). The use of social media at work for communications among healthcare professionals is increasing because it can provide instantaneous communications irrespective of the geographic locations of the two providers. Clinicians are texting and emailing each other to discuss cases (Kolowitz et al. 2014). Nurses are using social media at work to research the side effects of medicines and the symptoms of diseases. Patients are texting and emailing their doctors. Healthcare organizations like the Mayo Clinic are active in their use of YouTube, Facebook, and blogs. The goal of the Mayo Clinic's social media network is to improve global health by engaging patients using social media to share health information.

Social media use in healthcare has many benefits. The sender and receiver of social media information can quickly and easily locate each other, which saves time and reduces frustration. Collaboration is easier with social media because everyone can be online at the same time. With patients, social media has been shown to improve the health literacy skills of older adults, by increasing their cognitive functions through increased internet use and their engagement in social activities (Kobayashi et al. 2015), and to provide much needed healthy living information for rural adolescent mothers (Logsdon et al. 2015). State public health agencies are using Twitter and Facebook to disseminate COVID-19 statistics and information.

A few examples of ethical dilemmas in social networking include a family member who shared health information on an individual on their Facebook page, nursing students posting pictures of patients they come in contact with as a part of their educational training, and others posting pictures of individuals seen at a particular healthcare facility, such as a person leaving a Planned Parenthood clinic or substance abuse clinic. In many of these instances, this would not be against the law but would be unethical and likely against institutional policy.

Case Study: Social Networking Site Policies

Scenario: You are an HIM professional employed at a cancer clinic, and also a member of a private group hosted on a social networking site. Mary is not a member of the social networking group, but she is a friendly acquaintance of all the group members. Mary is a cancer survivor. She has attended many parties and community events with various members of the group. At these events, she has been very open about her cancer, her treatment, and her prognosis. It has been several months since the group members have seen Mary. One day Mary comes into the cancer clinic for a follow-up appointment while you are working. She encourages you to let "the gang" know that she is doing well. The easiest way for you to convey this great update is via the group's private site. What would you do?

Option 1: Go ahead and post the happy update to the group page; after all, Mary said it was okay.

Option 2: Do not put yourself at risk by using the social networking site at all. Instead, tell just a few of Mary's friends—the ones you are certain she would want you to tell.

Option 3: Explain to Mary that because you are learning this information about her in your capacity as an employee, it is important that any information sharing come from her—and not from you. You use this opportunity to talk to Mary about your organization's commitment to patient privacy.

Which option is the most ethical choice? Why?

Health Literacy

For patients to give informed consent for the use of their healthcare data, they must have the details of the consent form explained to them and they must understand those details. One factor that affects their ability to understand these details is health literacy, which is the ability to obtain, process, and understand basic health information (Center for Health

Care Strategies 2011, 13). Health literacy levels have been shown to be lower for individuals for whom English was a second language and among those living in poverty (Center for Health Care Strategies 2011, 13; HHS 2010). A patient with low health literacy lacks the skills to interpret medical documents and numbers, which affects their clinical health outcomes and increases healthcare costs (Chen et al. 2013).

The Department of Health and Human Services developed an action plan to study health literacy to guarantee that all US citizens have been provided with accurate, understandable healthcare information (HHS 2010). The plan was written to:

- Provide access to accurate and actionable health information
- Deliver person-centered health information and services
- Support lifelong learning and skills for good health

The seven goals of the plan are:

1. Develop and disseminate accurate, accessible, actionable health and safety information
2. Promote changes in the healthcare system that improve the health information, communication, informed decision making, and access to health services
3. Incorporate accurate, standards-based, and developmentally appropriate health and science information and curricula in child care and education through the university level
4. Support and expand local efforts to provide adult education, English language instruction, and culturally and linguistically appropriate health information services in the community
5. Build partnerships, develop guidance, and change politics
6. Increase basic research and the development, implementation, and evaluation of practices and interventions to improve health literacy
7. Increase dissemination and use of evidence-based health literacy practices and interventions. (HHS 2010, 6)

Specific actions include clear labeling of food and drugs, use of the internet and social media for disseminating health information, use of consumers to evaluate healthcare materials, conducting ongoing health literacy training at work and in schools, and providing patient-centered care (HHS 2010). Additionally, mandatory continuing education in health literacy for health professionals was suggested.

As of 2022, the Centers for Disease Control and Prevention (CDC) Health Literacy Action Plan is still in place. Per the CDC, "we need this plan because limited health literacy is a national public health issue that affects almost 9 out of 10 adults" (CDC 2022). The plan is used as a framework for making small incremental steps toward improving the accessibility of health information and ensuring the information is easily understood by all audiences (CDC 2022).

Check Your Understanding 13.2

Answer the following questions.

1. Which of the following is an issue with the electronic provisioning of healthcare?
   a. Loss of connection between patient and physician
   b. Faster access to patient data

c. Legibility of documentation
d. Appropriate authorization required for accessing system

2. The concept of autonomy in healthcare is most closely related to the concept of
   a. Security
   b. Privacy
   c. Trust
   d. Responsiveness
3. Which of the following is an HIT factor that can contribute to errors in care?
   a. Lack of technology skills
   b. Workflow and communication
   c. Limited access to clinical information
   d. Time to access needed documentation

4. Type(s) of personal health records are:
   a. Stand-alone
   b. Tethered
   c. Patient portal
   d. a and b
5. An individual who reports a wrongdoing is referred to as a
   a. Hacker
   b. Whistleblower
   c. Informatics specialist
6. The process of disclosing protected health information from health records is called
   a. Release of information
   b. Health information exchange
   c. Health literacy

Ethical Decision-Making

Handling ethical issues correctly falls into the realm of risk management. Conflicts resulting from differing ethical viewpoints can disrupt the workflow, cause upheaval between co-workers, and result in harm to patients and to staff morale. A plan for ethical decision-making should be put in place and employees should be given examples of correct and incorrect use of the process. Steps to manage ethical situations should include assessing all situations for their ethical subcontexts, identifying the underlying ethical issues, evaluating the situation including the perspectives of all involved in the situation, and deciding on a course of action. After the event, ongoing monitoring should be implemented to prevent future occurrences of similar situations.

Assessing Potential

This chapter has discussed many nuances of ethical issues related to technology that could form the basis for a risk assessment of potential ethical hotspots related to health informatics. Policies, procedures, hardware, software, and education surrounding the use of PHI, PHRs, HIEs, mobile devices, social media outlets, and release of IT must be assessed for their potential to cause harm to the patient.

Identifying

The PHRs, HIEs, mobile devices, servers, personal computers, and other computing systems should be catalogued and the level of risk potential for these systems should be rated. The potential of the human resources for generating unethical behavior should be added to this list. Several instruments are available for ethics screenings of employees.

For example, oncology and ICU nurses at Mayo Clinic attended an ethics workshop where they were trained to use an ethics screening tool that was developed by nurses to identify early warning signs of ethical conflict (Pavlish et al. 2015). Participants then evaluated the tool for three months. Results indicated the tool was beneficial in early identification, and the researchers recommend it be administered as part of the admission assessment.

Evaluating

Each item on the potential risk list should be evaluated step by step to determine when, where, and how the risk might occur. The cost and benefits of each risk should be evaluated. For example, while mobile application use is correlated with higher risks of data breaches, the use of mobile devices for looking up drug interactions or treatment modalities can provide better patient outcomes because online references are updated more frequently than paper-based medical references (Bullock 2014). The completed risk evaluation should be documented and put in a repository that is accessible to HI and HIM project managers as it will serve to assist them in their project planning with regard to costs, resources, and timelines.

Deciding on a Course of Action

For each item on the list of risks, several potential courses of action should be outlined. Using the weightings for each path of action, a cost–benefit analysis should be conducted to determine the critical path and decide on which of many paths is the most optimal one. These decisions could expose a need for creating or modifying the organization's policies and procedures. Once determined, the course of action should be carried out and the results documented.

Preventing Future Occurrences

Ongoing vigilant prevention is essential for the success of any risk management initiative. Indicators should be put in place and the necessary data collected for reporting on the previously identified risk situations. Mandatory annual training on ethics should be conducted for all employees. Top management should attend workshops and retreats designed to emphasize their role in creating an ethical work culture and rewarding ethical behavior.

Educational programs and practice in working through ethical situations can take many forms and can be adapted for different groups and different situations. One example of a tool to use in working through the ethical decision-making process is the matrix shown in Figure 13.4.

Figure 13.4. Ethical decision-making matrix

Steps	Information	
1. What is the question?		
2. What is my "gut" reaction?		
3. What are the facts?	Known	To be gathered
4. What are the values? \	Patient: HIM professional(s): Healthcare professional(s): Administrators: Society: Others as appropriate:	
5. What are my options?		
6. What should I do?		
7. What justifies my choice?	Justified	Not justified
8. How can I prevent this problem?		

Ethical decision-making is a process to be worked through with many aspects to consider. Solving ethical situations is not a quick decision to be made. Working through the steps of this process, the first thing to determine is the question. What is the question in a situation where a potential employer has been given genetic information that suggests the likelihood of a potential job applicant getting a disease that could be costly for the employer's health plan and could cause missed work if the employee were to actually develop the condition? Is the problem that the employer's livelihood and financial well-being could be jeopardized? Or is the problem that using this information in making an employment decision is considered discrimination? Next is the "gut reaction," which is setting aside consequences and personal feelings about the course of action to be taken. This should then be followed up with gathering pertinent information. In the employment issue, this could be finding out exactly what the law says about using genetic information in the hiring decision and working with human resources to assure the screening and hiring process follows procedures (Harman and Cornelius 2017).

The next step is to evaluate the values of all those involved in the particular situation. In the previous employment example, the values of the potential employee, the hiring supervisor, the organization, and others are in conflict much of the time. From all of this, options can be identified, followed by choosing what should be done. This should then be followed by articulating the justification for action to be taken based on the steps in the ethical decision-making. To avoid similar situations in the future, this final step is important for determining what could have been done to avoid the situation. This could result in rewriting policies and procedures, more training, or other similar activities. Although every situation will be different, using this process should lead to better decision-making.

Chapter 13 Review Exercises

Match the term with its description.

a. Electronic record maintained by patients for collecting and sharing their health data

b. Principles of right conduct that an individual adopts

c. What ethics and morals are based on
d. An individual's capacity to understand health information
e. Doing good for others

1. _____ PHR
2. _____ Values
3. _____ Morals
4. _____ Health literacy
5. _____ Beneficence

Discussion

6. As the HIM manager, what would be the best response to a complaint that one of your employees was posting information about a patient on Facebook?
7. How would you apply the AHIMA Code of Ethics to a report that a medical coder has consistently been upcoding medical records to increase reimbursement?
8. Apply the steps of ethical decision-making to a case where a physician is accepting money from a vendor to recommend the vendor's new EHR system.

REFERENCES

Abdelhak, M., S. Grostick, and M.A. Hanken. 2011. ***Health Information: Management of a Strategic Resource***, 4th ed. St. Louis, MO: Saunders.

AHIMA. 2013. "Redisclosure of patient health information (2013 update)." https://bok.ahima.org/doc?oid=300263#.YvELaLbMKCp.

AHIMA. 2016. "AHIMA standards of ethical coding." http://bok.ahima.org/CodingStandards#.YkmedonMKUk.

AHIMA. 2017. ***Pocket Glossary of Health Information Management and Technology***, 5th ed. Chicago, IL: AHIMA Press.

AHIMA. 2020. ***AHIMA Ethical Standards for Clinical Documentation Integrity (CDI) Professionals***. https://www.ahima.org/media/r2gmhlop/ethical-standards-for-clinical-documentation-integrity-cdi-professionals-2020.pdf?oid=301868 .

AHIMA. 2019. "AHIMA code of ethics." http://bok.ahima.org/doc?oid=105098#.YkmeOInMKUk.

AHIMA Professional Ethics Committee. 2013. Reality check 2013: Ethical issues in HIM. ***Journal of AHIMA***: web extra. http://journal.ahima.org/2013/10/01/reality-check-2013-ethical-issues-in-him/.

AMIA. 2016. "Ethical, legal, and social issues." https://www.amia.org/programs/working-groups/ethical-legal-social-issues.

Anderson, J. G. 2004. The role of ethics in information technology decisions: A case-based approach to biomedical informatics education. ***International Journal of Medical Informatics*** 73(2):145–150. doi:10.1016/j.ijmedinf.2003.11.015.

Budinger, T. F., and M. D. Budinger. 2006. ***Ethics of Emerging Technologies: Scientific Facts and Moral Challenges***. Hoboken, NJ: Wiley.

Bullock, A. 2014. Does technology help doctors to access, use and share knowledge? ***Medical Education*** 48(1):28–33. doi:10.1111/medu.12378.

Butler, M. 2015. Cracking encryption: Despite benefits, technology still not widely used to combat multi-million dollar breaches. ***Journal of AHIMA*** 86(4):18–23.

Cellucci, L. W., E. J. Layman, E. M. Campbell, and Z. Xiamong. 2011. Integrating healthcare ethical issues into IS education. ***Journal of Information Systems Education*** 22(3):215–224.

Center for Health Care Strategies. 2011. Health literacy interventions and outcomes: An updated systematic review. https://effectivehealthcare.ahrq.gov/products/health-literacy/research.

Centers for Disease Control and Prevention (CDC). 2022. "CDC's health literacy action plan." https://www.cdc.gov/healthliteracy/planact/cdcplan.html.

Chen, J.-Z., H.-C. Hsu, H.-J. Tung, and L.-Y. Pan. 2013. Effects of health literacy to self-efficacy and preventive care utilization among older adults. ***Geriatrics & Gerontology International*** 13(1):70–76. doi:10.1111/j.1447-0594.2012.00862.x.

Cheshire, W. P. 2014. Can electronic medical records make physicians more ethical? ***Grey Matters*** 30(3):135–141.

Clark, J. R. 2014. Making your mark. ***Air Medical Journal*** 33(5):194–196. doi:http://dx.doi.org/10.1016/j.amj.2014.06.007.

Cochran, G. L., L. Lander, M. Morien, D. E. Lomelin, C. Reker, and D. G. Klepser. 2015. Consumer opinions of health information exchange, e-prescribing, and personal health records. ***Perspectives in Health Information Management*** (Fall):1–12.

Crawford, M. 2011. Everyday ethics. ***Journal of AHIMA*** 82(4):30–33.

Cross, T. 2006. Academic freedom and the hacker ethic. ***Communications of the ACM*** 49(6):37–40. doi:10.1145/1132469.1132498.

Crysts, A. 2016. Preventing malpractice lawsuits due to EHR errors. ***Medical Economics*** 93(9):13–17.

Denić, N., V. Moračanin, M. Milić, and Z. Nešić. 2014. Risk management in information system projects. ***Upravljanja rizicima projekta informacijskih sustava*** 21(6):1239–1242.

Department of Health and Human Services (HHS). 2009. The Genetic Information Nondiscrimination Act of 2008. Information for researchers and health care professionals. Genetic Information Nondiscrimination Act Guidance (2009) | HHS.gov

Department of Health and Human Services (HHS). 2010. "National action plan to improve health literacy." https://health.gov/communication/initiatives/health-literacy-action-plan.asp.

Department of Health and Human Services (HHS). 2014. "Annual report to Congress on breaches of unsecured protected health information for calendar years 2011 and 2012." https://www.hhs.gov/guidance/sites/default/files/hhs-guidance-documents/breachreport2011-2012.pdf.

Department of Health and Human Services (HHS). 2015. "Breaches affecting 500 or more individuals." https://www.hhs.gov/hipaa/for-professionals/breach-notification/index.html.

Department of Health and Human Services. 2016. "Health information privacy. All case examples." http://www.hhs.gov/hipaa/for-professionals/compliance-enforcement/examples/all-cases/index.html#case15 .

Department of Health and Human Services, Centers for Medicare and Medicaid Services (CMS). 2012. Medicare and Medicaid programs: Electronic health record incentive program—Stage 2. ***Federal Register 77***(171):53967–54162.

Ellison, R., and L. Clack. 2021. ***Health Policy Management: A Case Study Approach.*** Burlington, MA: Jones & Bartlett Learning.

Han, Y., J. Carcillo, S. Venkataraman, R. Clark, R. Watson, T. Nguyen, H. Bayir, and R. Orr. 2005. Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system. ***Pediatrics*** 116(6):1506–1512.

Harman, L. B., and F. Cornelius. 2017. ***Ethical Health Informatics.*** Burlington, MA: Jones and Bartlett.

Health Insurance Portability and Accountability Act of 1996. Public Law 104-191.

HealthIT. 2016. "Are there different types of personal health records (PHRs)?" https://www.healthit.gov/providers-professionals/faqs/are-there-different-types-personal-health-records-phrs .

HHS. 2017. Health Information Privacy: All Case Examples. All Case Examples | HHS.gov

HIMSS. 2011. "Security of mobile computing devices in the healthcare environment." http://www.himss.org/security-mobile-computing-devices-healthcare-environment.

HIMSS. 2016. "Ethical issues in the use of health information." http://www.himss.org/ethical-issues-use-health-it.

Househ, M. S., E. M. Borycki, W. M. Rohrer, and A. W. Kushniruk. 2014. Developing a framework for meaningful use of personal health records (PHRs). ***Health Policy and Technology*** 3(4):272–280. doi: 10.1016/j.hlpt.2014.08.009.

Imgraben, J., A. Engelbrecht, and K.-K. R. Choo. 2014. ***Always Connected, But Are Smart Mobile Users Getting More Security Savvy? A Survey of Smart Mobile Device Users.*** Rochester, NY: Social Science Research Network.

International Medical Informatics Association. 2016. "The IMIA code of ethics for health information professionals." http://www.imia-medinfo.org/new2/pubdocs/Ethics_Eng.pdf.

Kirkpatrick, G. 2002. The hacker ethic and the spirit of the Information Age. ***Max Weber Studies*** 2(2):163.

Kobayashi, L. C., J. Wardle, and C. von Wagner. 2015. Internet use, social engagement and health literacy decline during ageing in a longitudinal cohort of older English adults. ***Journal of Epidemiology & Community Health*** 69(3):278–283.

Kolowitz, B., G. Lauro, J. Venturella, V. Georgiev, M. Barone, C. Deible, and R. Shrestha. 2014. Clinical social networking—A new revolution in provider communication and delivery of clinical information across providers of care? ***Journal of Digital Imaging*** 27(2):192–199.

Layman, E. J. 2008. Ethical issues and the electronic health record. ***The Health Care Manager*** 27(2):165–176.

Lefkowitz, J. 2006. The constancy of ethics amidst the changing world of work. ***Human Resource Management Review*** 16(2):245–268.

Levine, B. A. 2015. Digital OB/GYN: Gadgets for Health Tracking. ***Contemporary OB/GYN.*** http://contemporaryobgyn.modernmedicine.com/contemporary-obgyn/news/digital-obgyn-gadgets-health-tracking .

Logsdon, M. C., M. Mittelberg, and J. Myers. 2015. Use of social media and internet to obtain health information by rural adolescent mothers. ***Applied Nursing Research*** 28(1):55–56.

McWay, D. C. 2021. ***Legal and Ethical Aspects of Health Information Management***, 5th ed. Clifton Park, NY: Cengage Learning.

Mercuri, J. 2010. The ethics of electronic health records. ***Clinical Correlations: The NYU Langone Online Journal of Medicine.*** http://www.clinicalcorrelations.org/?p=2211.

Murphy, T., and A. Lau. 2008. Dealing with outliers. ***Datapoints.*** http://www.astm.org/SNEWS/ND_2008/datapoints_nd08.html.

NSW Health. 2016. "Fact Sheet 19/Ethical issues in human genetics and genomics." ***Health Centre for Genetics Education.*** http://www.genetics.edu.au/Publications-and-Resources/Genetics-Fact-Sheets/FactSheetELSI.

Olden, P. 2019. ***Management of Healthcare Organizations: An Introduction***. Chicago, IL: Health Administration Press.

Oxford Dictionaries. 2016. https://en.oxforddictionaries.com/.

Pavlish, C. L., J. Henriksen Hellyer, K. Brown-Saltzman, A. G. Miers, and K. Squire. 2015. Screening situations for risk of ethical conflicts: A pilot study. ***American Journal of Critical Care*** 24(3):248–257. doi: 10.4037/ajcc2015418 . https://depts.washington.edu/bioethx/topics/ethics.html

Pearlman, R. 2013. ***Ethics Committees, Programs and Consultation.*** Seattle: University of Washington School of Medicine.

Petersen, C., E. S. Berner, P. J. Embi, K. F. Hollis, K. W. Goodman, R. Koppel, C. U. Lehmann, H. Lehmann, S. A. Maulden, K. A. McGregor, A. Solomonides, V. Subbian, E. Terrazas, and P. Winkelstein. 2018. AMIA's Code of Professional and Ethical Conduct 2018. ***Journal of the American Medical Informatics Association*** 25(11):1579–1582. https://doi.org/10/1093/jamia/ocy092.

Statista. 2022. "Most popular social networks worldwide as of January 2022, ranked by number of monthly active users." https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/.

Studeny, J., and A. Coustasse. 2014. Personal health records: Is rapid adoption hindering interoperability? ***Perspectives in Health Information Management*** 11 (Summer).

Tan, M., and K. S. Aguilar. 2013. An investigation of students' perception of Bluetooth security. ***Information Management and Computer Security*** 20(5):364–381. doi: 10.1108/09685221211286539.

Tang, P., J. Ash, D. Bates, J. Overhage, and D. Sands. 2006. Personal health records: Definitions, benefits, and strategies for overcoming barriers to adoption. ***Journal of the American Medical Informatics Association*** 13(2):121–126. doi:10.1197/jamia.M2025.

Thomas, G. 2007. Secure mobile device use in healthcare guidance from HIPAA and ISO17799. ***Information Systems Management*** 24(4):333–342. doi: 10.1080/10580530701586060.

Vaas, L. 2007. Inside the mind of a hacker. ***eWeek*** 24(24):38–46.

van der Rijt, R., and S. Hoffman. 2015. Ethical considerations of clinical photography in an area of emerging technology and smartphones. ***Journal of Medical Ethics*** 40:211–212. doi: 10.1136/medethics-2013-101479.

Vest, J. R., and L. D. Gamm. 2010. Health information exchange: Persistent challenges and new strategies. ***Journal of the American Medical Informatics Association*** 17(3):288–294.

ADDITIONAL RESOURCES

AHIMA. 2022. "Ethics: Code of ethics case studies." https://www.ahima.org/who-we-are/governance/ethics/.

Archer, N., U. Fevrier-Thomas, C. Lokker, K. A. McKibbon, and S. E. Straus. 2011. Personal health records: A scoping review. ***Journal of the American Medical Informatics Association*** 18(4):515–522. doi:10.1136/amiajnl-2011-000105.

Aulisia, M. 2016. Why did hospital ethics committees emerge in the U.S.? ***AMA Journal of Ethics*** 18(5):546–553.

Bagley, J. E., D. D. DiGiacinto, and K. Hargraves. 2014. Imaging professionals' views of social media and its implications. ***Radiologic Technology*** 85(4):377–389.

Bates, M., C. Black, F. Blair, L. Davis, S. Ingram, D. Lane, A. McElderry, B. Peagler, J. Pickett, C. Plettenberg, and S. Hart-Hester. 2014. Perceptions of health information management educational and practice experiences. ***Perspectives in Health Information Management*** 11(Summer):2–12.

Bønes, E., P. Hasvold, E. Henriksen, and T. Strandenæs. 2007. Risk analysis of information security in a mobile instant messaging and presence system for healthcare. ***International Journal of Medical Informatics*** 76(9):677–687. doi:10.1016/j.ijmedinf.2006.06.002.

Buelow, J. R., P. Mahan, and A. W. Garrity. 2010. Ethical dilemmas as perceived by healthcare students with teaching implications. ***Journal of Teaching and Learning*** 7(2):85–92.

Darr, K. 2002. ***Ethics in Health Services Management***, 5th ed. Baltimore, MD: Health Professions Press.

Department of Health and Human Services, ONC, HealthIT. 2015. "Investigations of Health-IT related deaths, serious injuries or unsafe conditions. Final report." The Joint Commission. https://www.healthit.gov/sites/default/files/safer/pdfs/Investigations_HealthIT_related_SE_Report_033015.pdf .

Eddy, N. 2014. Mobile devices still unsecured in the workplace. ***eWeek***. http://www.eweek.com/mobile/mobile-devices-still-unsecured-in-the-workplace.html.

Health Information Technology for Economic and Clinical Health (HITECH) Act. 2009. https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/understanding/coveredentities/hitechact.pdf .

Kurzman, P. A. 2012. Workplace Ethics: Issues for Human Service Professionals in the New Millennium, In ***Encyclopedia of Applied Ethics,*** 2nd ed., edited by R. Chadwick, 559–564. San Diego, CA: Academic Press.

Lau, L. K., B. Caracciolo, S. Roddenberry, and A. Scroggins. 2011. College students' perceptions of ethics. ***Journal of Academic and Business Ethics*** 5:1–13.

Markel Foundation. 2011. "PHR adoption on the rise." http://markle-stage.svr04.com/publications/1440-phr-adoption-rise.

Putrillo, R. B., and R. F. Doherty. 2010. ***Ethical Dimensions in the Health Professions***, 5th ed. St. Louis, MO: Saunders.

Samuel, H. W., and O. R. Zaiane. 2014. A repository of codes of ethics and technical standards in health informatics. ***Online Journal of Public Health Information*** 6(2):e189. doi: 10.5210/ojphi.v6i2.5484.

Sittig, D. F., and H. Singh. 2013. A red-flag-based approach to risk management of EHR-related safety concerns. ***Journal of Healthcare Risk Management: The***

Journal of the American Society for Healthcare Risk Management 33(2):21–26. doi:10.1002/jhrm.21123.

US Department of Labor Family and Medical Leave Act. 1993. "Family and Medical Leave Act." https://www.dol.gov/whd/fmla/.

US Food and Drug Administration. 2016. ***"Mobile medical devices."*** http://www.fda.gov/MedicalDevices.

Chapter

14

Technology for Digital Health: Telemedicine and Mobile Health Systems

By Andi Saptono, PhD

Learning Objectives

- Explain the scope of digital health from the perspective of the domain's evolution over time
- Detail key characteristics of digital health that makes it impactful
- Explore the perspectives and expectations of stakeholders in digital health
- Identify the technologies used in digital health, with a focus on telemedicine and mHealth
- Identify challenges in health and potential solutions that can be achieved through the use of digital health
- Detail the roadmap of the future for digital health

KEY TERMS

Asynchronous interactivity
Consumer health informatics (CHI)
Digital health
e-health
Ecological Momentary Assessment (EMA)
Ecological Momentary Intervention (EMI)
Food and Drug Administration
Health information
Information and communication technology (ICT)
Internet of Things (IoT)
Medical device
mHealth

Mobile technology
Remote patient monitoring
Real-time interactivity
Synchronous interactivity
Store-and-forward interactivity
Telehealth
Telemedicine
Videoconferencing
Wearables

The advancement of informatics in health has allowed a convergence of digital technologies with health and healthcare. The goal of digital health is to improve efficiency of health, improve effectiveness of health services, and make medicines more personalized and precise. In this chapter, we will explore the definition of digital health to understand the scope of the domain, explore the stakeholders' perspective to help guide our investigation of the need and design of our digital health solution, and view the challenges as well as the risk of digital health in implementation. Afterward, we will do a deep dive into the technology used in digital health to help us be more aware of the potential solutions available that can address the gap and challenges in our stakeholders' health objectives. By exploring these key topics, we will trace the growth of digital health; where it was before and where it will go in the future.

Digital Health: Evolution and Definition

The World Health Organization (WHO) defines digital health as a "discrete functionality of digital technology that is applied to achieve health objectives and is implemented within digital health applications and information and communication technology systems" (WHO 2019a). AHIMA's definition on digital health further detailed the health objectives for individual aim, specifically adding "to enhance healthcare delivery and make approaches to health and wellness more personalized and precise" (Crabb 2019). From this definition, several notable key points can be observed: (1) digital health is a technology-based approach; (2) digital health is developed to achieve specific health objectives; and (3) digital health is implemented within the health domain. The stakeholders for digital health are patients, healthcare providers, healthcare entity administrators, and health data scientists (WHO 2019b). In this crossroad between health and informatics, health data scientists' role becomes important as digital health uses information technology (IT) and data analysis is generated through the use of those technologies in healthcare.

The digital health concept evolves as adopters of **information and communication technology** (ICT) become more proficient in their use. For example, around 1947, ICT was adopted to allow and enhance access to radiology services from distance in (Bashshur et al. 2016). ICT was adopted around 1960 to enhance general healthcare services, giving way to the emergence of Telemedicine (Bashshur et al. 2016). As the implementation strategy and technology matures, domain specific information are delivered in conjunction with traditional use of ICT to communicate to expand specific care or service, such as the delivery of prescription and medication information to allow remote refill process in pharmacy (telepharmacy), more detailed daily condition of patient in nursing facility allowing clinician to understand patient's condition better and can provide better recommendation and guidance for nurses (telenursing), daily

training/rehabilitation monitoring of sessions in home allowing therapy to be conducted in higher frequency (telerehabilitation), providing domain-specific education (tele-education, tele-mentoring), or supporting the administration of healthcare entity. The concept e-Health emerged when other Health IT technologies are introduced into the domain, such as integration with EHR (Eysenbach 2001, Norman and Skinner 2006). Nowadays, the continuous evolution of communication technology allows delivery of health services anywhere anytime, via mobile platform, wearables, and Internet of Things devices (Kukafka 2019). Facilitated by these platforms, digital health emerges as the umbrella term for identifying innovations and solutions to healthcare delivered through the use of digital information and communication technology (Iyawa et al. 2016).

Key Characteristics of Digital Health

To understand better about digital health, we need to be aware of the key characteristics of digital health: 1) that digital health follows a holistic perspective of health, covering all spectrum of health, from prevention, treatment, as well as rehabilitation/recovery; 2) that because of the previous characteristic, digital health commonly involves multidisciplinary professions in health, including, but not limited to, providers from healthcare entities, health informaticists, and social science professionals; which may lead to 3) that digital health is a complex solution (WHO 2019a). Awareness of these key characteristics will allow us to develop a solution that can address the need of all the stakeholders with design that allow users to navigate through the complexity of digital health.

Digital health solutions are not intended to be substitutes of functioning health systems (Brall et al. 2019). Instead, digital health should be used to complement and enhance already in-place health services. For example, rather than eliminating an existing rehabilitation facility in rural areas, telerehabilitation should be implemented to enhance the capacity of the clinic to provide specialty services by connecting with expert therapists at metropolitan clinics. This approach may introduce two potential benefits: first, the immediate impact would be in the expansion of rural clinic's capacity to provide service and the rural clinic therapist's capacity to provide specialty service by learning from the interaction between patient with the expert therapist. Second, due to the natural need to exchange information between rural therapists and expert therapists (for example, patient's data generated at home or rural clinic need to be discussed with expert therapists), this particular digital health solution can be used as a platform to initiate health exchange which may accelerate health information exchange efforts (Lehne et al. 2019).

Digital health solutions also inherit the limitation of ICT being utilized, which may significantly reduce the service, care, or conditions that can be addressed. For example, one concern in delivering rehabilitation therapy remotely is the inability to catch the patient if they fall. Due to this limitation, health informaticists and practitioners of digital health need to understand which health challenges can realistically be addressed by technology before deploying or implementing any digital health solution.

Stakeholder Perspectives and Needs of Digital Health

Adoption of a new technology depends on whether the technology meet the needs of its users. Therefore, in digital health, management of stakeholder needs are crucial in the successful adoption of digital health. To understand the needs of stakeholders in digital health, we must first identify who the stakeholders of digital health are. WHO identified

four primary stakeholders of digital health with perspectives that can guide its implementation: healthcare clients and patients/, healthcare entity providers, healthcare entity administrators, and health data scientists (WHO 2019b).

Health and healthcare clients want digital health to be utilized in delivering personalized communication about their healthcare. While patients still accept information not specifically targeting them, they still want the information to be somewhat relevant to their context. For example, services that announce public health information would still be of interest if it is selectively targeting patients that have conditions related to that particular information. Clients also would like to use digital health to connect with their peers by sharing and learning others experience allows patients to find their community and reduce the feeling of isolation. Digital health can empower health and healthcare clients to take ownership of their own health by allowing personal health tracking and citizen-based reporting. Techniques such as information gathering or utilizing crowd sourcing can provide snapshots of trends in public health, which in turn can lead to faster decision-making process. The availability of information through the internet also drives patients' needs in being able to find health-related information ready to be accessed at their fingertips. Similarly, the adoption of access to financial services through communication technology drives patients' needs to be easily completed by any health-related financial transactions integrated as part of digital health solution.

Healthcare providers look at digital health as a solution to increase their effectiveness and efficiency in delivering. Providers want digital health to be able to provide meaningful information about patients with minimal efforts to use as possible to minimize the burden of information for healthcare provider as well as minimize distraction from providing healthcare services. Any solutions that increase the burden (such as due to information overload) and effort required (such as solutions with usability or accessibility issues) are perceived as barriers to care. Providers also would like digital health to be smart, concise, and prompt in delivering the information that can facilitate informed decision-making during episodes of care. No single provider claims responsibility for managing a patient's episode of care from start to finish, thus ICTs are viewed as a platform that can facilitate communication with patients as well as peers, allowing access to experts where the availability of the needed expertise is otherwise limited (for example, when care is being continued and transferred from outpatient facility in metropolitan hospital to a rural clinic near the patient's home). Additionally, providers that experience professional isolation due to geographical location are looking into digital health to provide solutions in continuing education as well, to ensure that their knowledge on care is up-to-date and on-par with their peers; and that their care is supported by the latest, best available evidence. Digital health solutions also can provide solutions to improve efficiency in the coordination of care, including integration with other systems across the entity, such as prescription and medication management, as well as laboratory, diagnostics, and imaging systems (Crabb 2019).

Administrators and managers of healthcare entities want digital health to provide solutions in managing the operations of healthcare entities. When integrated as part of digital health solution, key management functionalities including human resource management, supply chain management, asset or facility management, and health financing would potentially improve the efficiency of management process. Common communication features, such as event notifications and messaging systems that allow them to disseminate information on a larger scope, can be used during a crisis. Administrators and managers of healthcare entities also want digital health solutions to assist monitor and ensure the survivability of the healthcare entity (WHO 2019b).

Data scientist is the latest role to be introduced at the table of healthcare. Traditionally, data scientists are clumped together with administrators and managers and perform their roles in the background of the health entity's operation. With the emergence of digital health, data scientists have found themselves pushed into an active role with the responsibility to become the steward of patients' care. In this role, data scientists want digital health solutions to be the platform that allows them to collect high-quality, high-fidelity data from both the patients and the providers. Advanced algorithm, such as artificial intelligence, would power automatization of processes within digital health, especially to improve data quality by coding and parsing of unstructured data. Data scientists also play an important role in ensuring that patients' data can flow smoothly across entities to facilitate continuous care across the health spectrum.

Professional Organizations' and Health Entities' Perspectives on Digital Health

Professional organizations have recognized the emergence of digital health and have published their perspectives on topics such as telehealth and mobile health, better known as **mHealth**. For example, the American Medical Association (AMA) has published several "playbooks" in their Digital Health Implementation Playbook Series, including Playbook on Telehealth Implementation, Remote Patient Monitoring, as well as a playbook to deliver education for clinicians and organizations to better prepare them in implementing telehealth (AMA 2018). These playbooks presented a detailed strategy on how to approach a digital health solution and implement it within a healthcare entity. For example, the Telehealth Implementation Playbook focuses on introducing telehealth concepts and how they can affect practice, while also presenting content that helps to prepare healthcare entities to plan for their telehealth "pregame" (detailing strategies to identify need, form teams, define success via outcome, build the case for telehealth, and connect with vendors and technology providers), during "game time" (presenting strategies to develop workflow, form the day-to-day care team, implement technology and solutions, evaluate the solutions, and scale the solutions when needed), and the "postgame" period, including methods to perform follow-up and measure outcomes (AMA 2021).

page 402

To increase awareness of digital health, professional organizations such as the American Occupational Therapy Association, American Speech Language and Hearing Association, and American Physical Therapy Association have also joined the domain by providing continuing education and certification processes surrounding topics of interest in digital health, including regulation, administrative, technical, clinical, and ethical aspects of digital health. Meanwhile, the American Telemedicine Association (ATA) has been leading the charge in telehealth adoption by promoting policy changes in regard to reimbursement of services and licensure (ATA 2022).

Challenges in Adopting Digital Health

Digital health faces several key issues that need to be addressed before the solution can be adopted as part of approaches to health and wellness, including risks to security, privacy, and confidentiality of protected health information (PHI) due to the use of public communication infrastructure, licensure and credentialing of services, lack of evidence that can lead to limited reimbursement for services, and ethical issues around digital health (Kostkova 2015).

Although advanced IT strategies (such as stronger encryption, multifactor authentication, and access role restriction complemented with audit trails) have been introduced to mitigate some of the security risks association with digital health, the use of public communication infrastructure, such as the internet, still opens the potential for a security breach of PHI. Phishing, malware, and, most recently, smishing (attempt to breach security via the use of SMS and text messaging) are still used to lure and trick unsuspecting individuals into permitting access to their system (Loumbas and Peters 2022). Not only do these security risks open up a path for bad actors to acquire access to data, but healthcare entities find it difficult to integrate patient-generated data into their system due to potential error or reduced data integrity caused by hacking. Security systems powered by artificial intelligence might help mitigate newer types of attacks that mimic human behavior by identifying and marking suspicious messages and attempts to breach security protocols (Wahsheh and Al-Zahrani 2021).

The rapid growth of technology has driven the growth of digital health. New, faster computing systems are introduced every year. Unfortunately, this very situation makes it challenging to develop digital health solutions due to the rapid introduction of newer solutions that aim to deliver similar service with newer technology innovations (Guo et al. 2020). Outcomes measured from implementing a digital health solution might not be generalized to newer systems due to the rapid evolution of the technology that might already have addressed the limitations of the previous system. Evidence is scarce and usually produced from study with small number of participants or scenarios, which creates reluctance in adopting them to change policies. Without changes in policy, especially in reimbursement, digital health faces challenges to its sustainability as many systems did not survive beyond the initial research or development funding that was allocated at the beginning of the project.

The situation mentioned in paragraph above, however, might change soon due to the emergency health status caused by the COVID-19 pandemic. The pandemic relaxed many healthcare-related laws and resulted in funding for many digital health initiatives and Medicare coverage of digital health efforts, such as telehealth and remote patient monitoring. Private insurers have followed this example by creating waivers to reimburse for services delivered using digital health. This situation allows many outcomes of digital health to be measured, and the resulting data could potentially be aggregated into evidence that could be used to affect policy beyond the pandemic (Gunasekeran et al. 2021). In a sense, the pandemic forced decades of regulatory change crammed into roughly a year and a half. Although many of the relaxed laws might be discontinued, the knowledge and evidence gathered during this period could lead a future where digital health can grow even further and be adopted as part of the day-to-day health service offered by healthcare entities.

Common sentiment on digital health pointed that patients prefer the convenience offered by digital health solutions and will demand it as part of their options moving forward (Jetty et al. 2021). Strengthening this sentiment, these same publications also highlighted that certain area of healthcare reported that their services can be delivered effectively using telehealth (for example, about 25 to 30 percent of primary care visits).

Introduction of new technology and innovations in digital health also introduces new ethical challenges in the delivery of care. Following the traditional approach of medical ethics, four key ethical issues need to guide the delivery of care: autonomy, beneficence, nonmaleficence, and justice (Varkey 2021). Digital health practitioners need to be mindful that the implementation would enable patients to make their own health-related decisions. Digital health solutions should also be used only when the technologies have been shown to provide benefits while minimizing risk and harm to the patients. Introduction of services delivered via digital health should improve access to care instead

of limiting the patients' choice of services. For example, some reports identified that care for individuals with disabilities has become limited due to the introduction of the telerehabilitation service that forces them to switch to in-network service delivered remotely instead of continuing the out-of-network, in-person care that they were using prior to the introduction of telerehabilitation. This situation is contrary to the very goal of digital health (to not replace a working health system) as well as posing issues in the ethics of care.

Check Your Understanding 14.1

Answer the following questions.

1. Which of the following are not considered to be stakeholders of digital health based on the World Health Organization's definition?
   a. Patients
   b. Healthcare providers
   c. Healthcare administrators
   d. Data scientists
   e. Educators
2. What are the key characteristics of digital health?
   a. Temporal, ethical, universal
   b. Interconnectivity, asynchronous, continuous
   c. Complex, multidisciplinary, holistic
   d. Accessible, usable, personable
   e. Educative, evocative, entertaining
3. Which statement best captures one expectation of stakeholders in digital health?
   a. Patients are expecting to be able to deliver care themselves.
   b. Providers are expecting to have fewer patients due to digital health's automation.
   c. Administrators are expecting that digital health will market more services to more patients.
   d. Data scientists are expecting that digital health will allow them to collect better quality data about health.
   e. Educators are expecting that digital health will allow them to deliver more educational materials.
4. What are some commonly used techniques to mitigate security risks in digital health?
   a. Biometric and telemarket
   b. Self-check application and dongle
   c. Security by obscurity
   d. Role-based access and database access
   e. Encryption and multifactor authentication
5. Digital health should operate following the traditional approach of medical ethics, including:
   a. Automation, beneficence, nonmaleficence, and justification
   b. Supervision, justice, beneficence, and helpfulness
   c. Autonomy, beneficence, nonmaleficence, and justice
   d. Self-support, assistance, access, and equity
6. Rapid growth of technology drives advances in digital health but also introduces new challenges, such as:
   a. Security and privacy risks
   b. Social media crosscheck

c. Proliferating health news from multiple channels
d. Integration with personal records
e. New networks to provide service

7. Patients of digital health are expecting that:
   a. Medicine can be replaced by technology
   b. Communication and care will be personalized to their context
   c. Travel will no longer be required to access any service
   d. They will have freedom in choosing providers
   e. Healthcare entities provide the technology for accessing digital health
8. Digital health may facilitate:
   a. Faster medication orders
   b. Access to personal health history for organizations and institutions
   c. Spread of medical tourism
   d. Advanced technology such as universal medication box
   e. Interconnectivity between fractured systems in healthcare

Provide a short answer for the following.

9. Who published the Digital Health Implementation Playbook Series?
10. Identify a policy that positively affected digital health efforts during the COVID-19 pandemic.
11. Identify potential cyberattacks that a digital health solution might experience.

Technology in Digital Health

As mentioned earlier, digital health solutions are developed using the information and communication technology (ICT). Users of digital health may communicate, interact, as well as perform activities in real-time (synchronously) or asynchronously to achieve their health objectives. When determining a digital health solution to implement, we need to identify what type of interactivity that our users will need, understand the pros and cons of each type of interactivity, and choose the appropriate technology that would provide the type of interactivity needed. Therefore, in this chapter, our first step is to study the two forms of interactivity for digital health: synchronous and asynchronous.

Synchronous Interactivity

Synchronous interactivity, also known as real-time interactivity, is the form of communication between users of digital health where the communication channel allows transmission of information in both directions (incoming and outgoing) at the same time. This type of communication channel allows information to flow at any time from any point of the communication to their intended target. Because of the characteristic of the channel (where information can flow from any participant at any time), any participants in the communication channel can produce and receive responses in real time.

Synchronous interactivity is often implemented through the use of real-time communication technology to connect its participants, such as **videoconferencing** and audio communication. An example of this is when a physician has a teleconsultation with a patient via videoconferencing. Through this means of communication, providers can interact directly with their patients to analyze their condition, with patients providing information that can help the process of diagnosis. This communication channel allows high-intensity information exchange between participants, allowing inquiries, responses, and feedback to be given in real-time, which can be used during a consultation or an educational session. Multiple healthcare specialists can attend the same session to offer

insight to the situation at hand, which can facilitate smooth multidisciplinary collaboration to solve a particularly challenging situation. The real-time interactivity also nurtures relationships, with many patients feeling that their providers are "there" with them even if they are remote. However, communicating in real time also requires that all participants connect at the same time, which might limit the opportunities to communicate due to scheduling or availability constraints.

As mentioned earlier, synchronous interactivity usually allows high-intensity information exchange between its participants. Due to the high intensity of the interactivity, the synchronous form of communication often requires high-end computation, network, and information technology specifications that could have relatively high equipment costs, depending on how reliable the system needs to be. For example, in a tele-surgery where information about patient need to be streamed as accurately and as fast as possible, high-end computers, sensors, and high capacity network bandwidth would be required. The visual data need to be sent without delay as even a millisecond delay may have negative impact on the surgery process. In addition, backup network to ensure that the communication stays up and reliable need to be established as well. Compare the scenario with a telerehabilitation scenario where a clinician watches and provides feedback as a patient perform a physical exercise at home. In the second scenario, the clinician might only need a gross movement pattern of the patient to ensure that the patient is performing the exercise correctly. Therefore, the need of the computers as well as bandwidth to sufficiently implement the telerehabilitation solution may be lower compared to the first scenario.

Asynchronous Interactivity

Asynchronous interactivity, on the other hand, only allows the communication channel to send information in one direction at a time. Any response from participants of the communication has to wait until the communication channel is available before transmission of information can occur. For example, the text message from a patient need to arrive at the clinician's device before the clinician can respond to the message. Contrast this with a real-time communication, in which clinician may respond mid-way even before the patient completed their sentences. This form of communication is also called store-and-forward when responses, messages, or information transmitted in the communication are stored temporarily within the communication channel while waiting for it to be available before being pushed to its intended target (for example, sending text message from patient to clinician, or from wellness coordinator to clients). Technologies commonly used to provide this form of interactivity include email, SMS text messages, texting via an Internet-based application, application programming interfaces (APIs), web services, and online portals.

Patients often report that asynchronous interactivity creates a feeling of distance between them and their providers (Breton et al. 2021). However, in cases when the quality of the communication network allows information to be transferred in a fast and efficient manner, participants might feel that their communication is in real-time instead of an asynchronous interactivity. For example, a fast response by clinician during text messaging session might give the impression that the communication happens in 'real-time', even when the underlying technology is asynchronous. In this scenario, the illusion of 'real-time' would easily be broken when the response stops abruptly or is delayed.

Because of its storing data prior to forwarding them to their intended recipient, asynchronous forms of communication are most often used to acquire and store clinical information. Health information can be stored and subsequently forwarded ("pushed") or retrieved ("pulled") by a healthcare entity to follow a patient's continuum of care. When time is not a constraint, asynchronous communication allows large amounts of data and

health information (such as vital signs data, x-rays, or dermatology images) to be transferred from one site to another in small chunks over longer period of time. In this scenario, the transfer of data via asynchronous communication can be performed over simple technology with significantly lower cost of computing and smaller bandwidth requirement for the network technology.

Telemedicine and Telehealth

Telemedicine is commonly defined as the use of health and medical-related electronic information exchanged between remote sites using ICT to improve patients' health status, support care, education, and even beyond, reaching to topics related to public health and health administration (ATA 2018; Barbosa et al. 2021). AHIMA describes telemedicine as the professional services provided to a patient through an interactive telecommunication system by a practitioner at a distant site (AHIMA 2017). This idea of telemedicine has been envisioned for more than 50 years; for example, some health-related comics from the 1960s have presented ideas of medical information being transmitted via air (using radio technology), and when combined with medical sensing equipment, allowing healthcare practitioners to make clinical decision and prescribe treatment for the patient.

Telehealth is commonly defined as a telecommunications system that links healthcare organizations and patients from diverse geographic locations and transmits text and images for (medical) consultation and treatment (AHIMA 2017). Telehealth is commonly perceived as an evolution of telemedicine where the utilization of a variety of telecommunication tools, including telephones, smartphones, and mobile wireless devices, with or without a video connection, are geared toward supporting the complete continuity of healthcare (Dorsey and Topol 2020). The use of telehealth to deliver health services is expanding due to technological advances and cost pressures for delivery of services to the patient. Examples of telehealth services include monitoring and reporting of vital signs and oxygen levels, teleradiology with results transmitted electronically to another site for diagnosis, videoconferencing with physicians and nurse practitioners, and online eye exams.

The COVID-19 pandemic accelerated the digitization of the healthcare industry. Decades of regulatory change and consumer adoption were crammed into a few short months (Stokel-Walker 2020). The global market for telemedicine was estimated to be worth around $80 billion in 2020 (Forbes Business Insight, 2021). Virtual care visits were reported to increase more than 60 folds from 2019 to 2020 (HHS 2021). Frost and Sullivan market research company projected a 38 percent annual growth rate between 2020 and 2025, culminating to about seven-fold growth by 2025. An increasingly aged population, increasing incidences of chronic diseases, and a rapid rise in the software market are all contributing to the growth of the market (Dorsey and Topol 2020). Telehealth represents an attractive solution to increase the quality of care while reducing overall healthcare expenditures.

Telehealth has been identified to provide benefits, including improved access to information and delivery of care not previously feasible. For example, clinician can now provide on-the-job assistance for individuals as part of their return-to-work process. Traditionally, clinician will have to travel and accompany their clients on the job, which limit the amount of assistance provided. With Telehealth, clinicians can perform remote monitoring for their clients and provide support as well as feedback remotely. Access has been the main theme of telemedicine, with efforts focusing on three aspects: increasing the amount of service or care that can be accessed, improving the quality of access to the service, and increasing access to the technology used in telemedicine. Some beneficial side effects of Telehealth have been reported by healthcare providers as well, including the reduction of the feeling of professional isolation from healthcare providers in rural

areas, the opportunity to further improve professional education via consultation and mentoring with experts from different sites, and improved knowledge about clinical issues and solutions from exchange of information between providers.

Telehealth can employ both synchronous (for example, using videoconferencing) and asynchronous (for example, using patient portal) interactivity. The primary use of Telehealth's synchronous activity is for consultation. Prior to the COVID-19 pandemic, the US had low rates of telemedicine utilization, with high telehealth adoption sites performing fewer than 100 video visits per day. Between February and May 2020, telehealth virtual visits at UVA's Center for Telehealth reportedly grew by 9000 percent, with some Telehealth sites reporting up to 50,000 visits per week (Healthcaredive, 2020). The technology primarily used for synchronous Telehealth is videoconferencing. Asynchronous Telehealth efforts are commonly found in scenarios when a patient's information must be sent before a synchronous interaction, such as image transmission and **remote patient monitoring**. With the improvement of ICT (such as improved bandwidth, faster internet connections, new types of media and platforms to transfer data to, and ubiquitous use of smartphones), the cost of Telehealth can be reduced due to availability and increased use.

There are many uses for Telehealth when coupled with some innovations in predictive diagnosis and feature determination and analysis. For instance, retinal images can be taken with cell phones and transmitted to a large medical center where reviews enhanced by artificial intelligence can examine an image and make clinical decisions. This technique can be used for diagnosis in ophthalmology with items of high incidence, such as diabetic retinopathy, age-related macular degeneration, glaucoma, premature retinopathy, age-related or congenital cataracts, and retinal vein occlusion (Nuzzi et al. 2021). This process can also be used in dermatology and other disciplines where an image can be critical in making a diagnosis and then determining subsequent treatments. Telehealth makes this possible by giving access to poor communities. This type of interaction increased exponentially during the COVID-19 pandemic, with patients unable or unwilling to travel to a hospital or doctor's office out of concern for exposure to active viruses. Many of those patients sought in-home care via cell phones or Internet video chats as opposed to physically entering a hospital or doctor's office.

Despite the advantages telehealth offers, it also brings barriers and challenges, including the feeling of a compromised relationship between health professional and patient, also known as depersonalization. This particular challenge is commonly reported in older adults utilizing Telehealth that use store-and-forward technology, in which patients do not interact directly with their providers. On the other hand, younger patients who are more accustomed to technology do prefer virtual care, which has been shown to be able to be delivered in between 25 and 30 percent of primary care visit scenarios. Another challenge is the amount of information captured from multiple sources (such as wearable sensors, video or audio consultation, or EHR). The stream of information from these sources can introduce information overload to healthcare providers that could affect the clinical decision-making that reduces the quality of care being rendered. Finally, prior to the COVID-19 pandemic, fear of liability was also reported as a major barrier to adoption of Telehealth (Gajarawala and Pelkowski 2021). Although some key policies and laws were relaxed during the public health emergency (such as licensing requirements and interstate compacts as well as prescription of controlled substances via Telehealth), currently, it is still unclear on whether the changes would stay post pandemic. Only by addressing these challenges would Telehealth be adopted fully into healthcare which would ensure the sustainability of Telehealth in the future.

Case Study: Telehealth Solution for Complex Rehabilitation Technology

Complex rehabilitation technology (CRT) is a field that focuses on providing medically necessary, personalized seating and mobility solutions for individuals with mobility disability. CRT products differ significantly from durable medical equipment due to the need for evaluation, fitting, configuration, and adjustment, as well as programming in some of the solutions for individuals with disabilities to meet their specific and unique medical, physical, and functional needs. Because of this, services related to CRT require more knowledgeable, skilled, and experienced professionals trained to conduct specialized evaluations, measurements, and ongoing modifications. These skill sets are not easily found in rural clinics, thus requiring patients to travel long distances to meet their providers, who are usually located in metropolitan area clinics.

page 410

Telehealth has been proposed as a solution to provide access in these areas, especially during the COVID-19 pandemic. *Clinician's Guide to Use of Telehealth for CRT Service Provision* was published in 2021 by Clinician Task Force (Clinician Task Force 2021), a nonprofit organization comprised of professionals in seating and wheeled mobility. The publication outlines a workflow chart that helps clinicians to make decisions on when and why telehealth would be appropriate to connect with their patients. The publication also includes several case scenarios where telehealth is utilized for evaluation, fitting, and supervising modification of CRT. A videoconferencing system is the recommended technology to deliver services in these scenarios.

The use of telemedicine for CRT has been reported to provide benefits such as flexible scheduling; streamlining of services from the initial prescription of CRT to routine follow-ups, which reduces the number of overall encounters; minimizing client travel, risk of exposure to infection at clinic, and expenses associated with travel; and lending itself to screening and triaging patient need. Telemedicine allows clinicians to observe patients' natural living environment, delivery process, and follow-ups. This can help them to provide additional recommendations and prescribe home modifications as needed to allow patients to optimally use their CRT in daily activities. Prescription via telemedicine has reported similar, and often better outcomes, than in-person services with patients and providers reporting high levels of satisfaction.

Telemedicine should not replace all in-person visits, but it can be a tool for successful treatment interventions and helping patients reach their health goals in a more efficient manner. The Clinician's Guide presented a scenario when telemedicine is not appropriate to provide service remotely to a client. This scenario warns that telemedicine should not be a replacement for every in-person service, especially when the limitation of the need for physical assessment cannot be addressed. Instead, it should be a complement to enhance the service experience.

The practice of medicine and public health supported by mobile devices is known as mobile health. Mobile health, also known as **mHealth**, is a form of telehealth where wireless devices and phone technologies are used for the generation, aggregation, and dissemination of health information between patients and providers (AHIMA 2016a). mHealth can use both forms of interactivity, synchronous and asynchronous, but most mHealth goals can be achieved using asynchronous forms of interactivity. mHealth uses mobile platforms such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices. Healthcare providers can use mHealth to access health information, communicate with patients, monitor patients' health conditions remotely, or conduct health-related research. In turn, patients can use mHealth to track their own health, access health information, or communicate with providers. mHealth applications are among the top apps changing healthcare, from "seeing" a physician via video consultation, having prescriptions filled, finding specialists for very specific conditions, to maintaining personal health records (PHRs), including recording daily metrics, such as blood pressure, oxygen levels, blood sugar readings, and so on (***PC Magazine*** 2015). Other mHealth apps include prescription services by mail, with medications prepacked and dispensed as individual doses; services of a personal health coach for a variety of conditions; and those used in clinics for capturing assessment results in clinic for the provision of care. There are devices on the market now that can record a medical-grade EKG and give an instant analysis right on a user's phone (Al-Alusi et al. 2019).

The wide spectrum of mHealth applications and needs as described above drives the global market for mHealth. mHealth market was estimated to be around $46 billion in 2020, with predictions that the market could reach more than $200 billion by 2026. This is due to the estimated growth of 30 percent per year (compared to EHR's growth at around 4 percent per year; Fortune Business Insight 2021; Fact, MR 2022). The drivers for mHealth growth are the prevalence and penetration of smartphones and tablets and the growth of the elderly population who would like to age in place and can utilize mHealth to manage their own health and chronic conditions. In this population, healthcare providers also see the potential of mHealth to provide remote health monitoring to ensure the wellness of the elderly individual.

The typical patient on average visits their clinician three times a year with an average encounter of 18 minutes (Neprash et al. 2021). mHealth maintains the relationship of patients and providers outside these medical appointments using key features, such as apps, secure messaging, mobile platform-based wearables and sensors, notification and reminders, and as many asynchronous services that can be accessed via mobile devices. While most telemedicine goals involve delivering specific health services, mHealth's goal is to maintain the established relationship between patients and providers over a long period of time. Because of the nature of this goal, most mHealth uses asynchronous interactivity and can be delivered inexpensively. However, care should also be exercised in selecting mHealth applications to implement as support for care because most mHealth applications have not been validated with data and research conducted on their usefulness are still in initial phase due to the newness of the technology.

Mobile Technology for Health

People have been trying to use technology to improve awareness of health and deliver health services since the printing press was invented, and media was introduced as platform to deliver information to the mass. One reason for utilizing mass media is the ubiquity and accessibility of the media, mainly because the media can reach a large number of people (AHRQ 2022) . Technologies such as TV or radio fall into this

category. After the proliferation of personal computers and the internet, people saw the potential of the internet-based computing platform due to its interactivity, allowing patients and providers to communicate using personal computers and the Web. None of these technologies, though, have all the features of personalization and context sensitivity that are found with mobile technology. **Mobile technology** introduces context sensitivity, which means that features like location can be used to deliver different types of interventions to match the patients' situations and conditions as they live their daily lives.

One of the important features of mobile technology is captured in the "always on, always carry" idea, as people carry smartphones all the time. On many occasions, people spend more time with their smartphones than with their significant others. Because the smartphone is "always on and always carried", mobile technologies occupy the area that anthropologist Edward Hall called "intimate" space (Hall 1963). He divided interpersonal distances into four zones: public space, social, personal, and intimate space. The intimate space allows unprecedented potential for health intervention, as it allows healthcare to be delivered directly to individuals while they go about their daily lives. This can take many forms, such as reminding them to take medication or to perform self-care tasks. Mobile technology also allows patients to assess fleeting health events, like seasonal depression or pain, in real time. These events are the kind of important health phenomena that patients will not be able to recall accurately if they are being assessed a day or even a week after the event. In psychology, this process is called the ecological momentary assessment (EMA), meaning the ability to assess behavior when it is happening in real time. The delivery of an intervention to answer an event detected by EMA is called the ecological momentary intervention (EMI). With this approach, mobile technology enables healthcare providers to activate a therapy app to help patients with depression right then and there. It also allows context-sensitive health interventions. As an example, when a child is experiencing anxiety while waiting for a bus, he or she can activate an app that helps with anxiety.

mHealth Implementation

When implemented in a clinical setting, mHealth is primarily used to improve communication between a healthcare provider and patient while providing an efficient way to access health information. For example, a clinic could employ the use of a tablet computer to allow nurses to access EHR on the go. Smartphones can also be used to coordinate care and connect healthcare providers from anywhere at any time.

This convenience comes with challenges, such as security risks (a misplaced or stolen tablet might contain PHI) and infection control (bringing an unsterilized tablet to a high-risk ward might increase the risk of in-clinic infection). Other challenges in adoption of mHealth in clinical settings include the possibility of mHealth distracting providers during care, potential interference with medical equipment that shares similar bandwidth or wave range with mHealth solutions, and lack of quality control. As mHealth itself is considered new to the field of healthcare, studies investigating appropriate mHealth use are still in early phases. Without addressing these challenges, healthcare providers may find it difficult to adopt mHealth in daily care activities.

In the home setting, mHealth is primarily used for remote monitoring and personal health tracking. Apps used in this approach include coaching or education on health conditions, progress recording, and continuous but noninvasive monitoring. The data captured by the app can be transmitted to clinicians and can be processed to show visualization of trends that might help in clinical decision-making (for example, using line charts to show progress of patient's care). In this setting, **wearables** are commonly used to track personal health. Combining wearables and mHealth allows interventions that can introduce behavior change; patients can perform activities, get notifications, and

review their own progress, while healthcare providers can deliver recommendations to assist patients in their care. Healthcare providers can also use mHealth to deliver health-related educational materials to home setting, allowing patients to learn more about their condition and create strategies to adopt the care process in their daily activities. These scenarios show the potential of mHealth in home, especially for individuals that would like to age in place.

With the strong potential of mHealth use in home setting, regulations on mHealth are focused on applications that are considered to be part of a category called **medical devices** (FDA 2019). The **Food and Drug Administration (FDA)** policies focus on regulating mHealth solutions intended to be used as an accessory to a regulated medical device (whether it is still in the process of getting an approval or already approved), or solutions that transform a mobile device into a regulated medical device. For other types of mHealth solutions, the FDA currently intends to exercise enforcement discretion on solutions that pose minimal risk to patients and consumers.

With more powerful mobile devices that employ faster processors, greater storage, and faster connectivity, mHealth continues to expand and deliver more features to support and maintain the relationship between patients and providers. More precise and miniaturized sensors allow better sensing capability. Coupled with big data analytics and artificial intelligence, mHealth is able to deliver features that are more personalized, precise, and easier to understand and use, at the same time providing the convenience of being available to use anywhere at any time.

E-Health

E-health is "the field in the intersection of medical informatics, public health, and business, which refer to health services and information delivered or enhanced through the Internet and related technologies" (Eysenbach 2001). One facet of e-health is the use of the Internet to deliver health information that informs consumers and online portals that empower patients to check their labs and make appointments. E-health can be used to provide easily accessible educational modules as part of a public health series for health workers. E-health also benefits healthcare businesses that advertise and sell their products using the Internet (Mobisoft 2018). In essence, E-health term expands the concept of Telehealth to include teaching, research, and business in relation to public health.

E-health for individual patients provides tools such as resources to learn about a diagnosis and online portals to access physician office records, appointment scheduling, and reminders. The e-health concept is expanding Telehealth to improve the efficiency and effectiveness of public health management of our communities, which will ultimately positively affect the health of individuals within the community. Data are collected on health status and behaviors as well as socioeconomic, environmental, biomedical, and genetic influences. It is also amassed on outcomes, use of resources, financial aspects of costs, and expenditures. These data are stored in a wide variety of databases that include EHRs, registries, and required surveillance and reporting locations. The distribution of clinical data is highly fragmented due to the lack of willingness to share data between healthcare stakeholders (Naeem et al. 2022). Some of the factors identified includes lack of trust in patient generated data, lack of incentives to share data, and lack of awareness of the importance in sharing data for public health. Once these issues are addressed, the potential for e-health applications is significant with the ability to use this data, including real-time monitoring of public health situation, disease prevention, and reducing health disparities in populations, which would benefit not only the community but the individuals within as well (van Panhuis 2014).

Consumer health informatics (CHI) is the area of HI focusing on the interaction of consumers, patients, and others with health information systems and applications. Some popular terminologies to describe CHI include patient-focused informatics, health-literacy tools, PHR (as opposed to the EHR, which is usually focused on use by the healthcare entity), or consumer education. This field differs from telemedicine, which focuses on delivering care from provider to patient, and mHealth, which focuses on maintaining the relationship between provider and patient. Instead, CHI mostly focuses on providing health information and health information tools for patients' own personal use. It is devoted to empowering patients to take ownership of their health by further evolving health technology with seamless connectivity into the environment via the **Internet of Things (IoT)**. According to Oxford Dictionary, IoTs are computing devices embedded in everyday objects that are constantly connected to the Internet which allow them to send and receive data as necessary. The tools or systems in CHI can be used with the assistance of a healthcare professional but does not necessarily depend on a healthcare professional to operate. Because of this, CHI is devoted to developing HI solutions based on patients' views.

The availability of CHI has been dependent on evolving IT. Due to the nature of CHI, most technology used in this domain provides asynchronous interactivity between health stakeholders. The current technology-enabled tools making health information available to consumers include the following:

- Online self-help books for use on e-readers and computers
- Online health articles or sites
- Bulletin boards and online discussion groups
- Automated appointment reminders
- Automated systems for learning more about health conditions
- Wearable sensors, such as glucose monitors
- Social media sites such as Facebook and Twitter
- Smartphone apps
- IoT

There are a wide variety of interactive systems where health information consumers can communicate with others using email, bulletin boards, online discussion groups, and social media. The communication is not limited to healthcare professionals, but also includes self-help groups. The initial technologies might seem antiquated in comparison to the myriad of technologies that today's health information consumers can choose from. Increasingly sophisticated technologies are becoming available to deliver health information to patients and to their healthcare providers. A study on improving hospital patient recovery time assessed patients who had elective cardiac surgery and who were provided post-operatively with a Fitbit to record the time when they began walking and the number of steps that they took each day. The results of this study revealed that the earlier in their recovery and the more walking they did, the shorter their length of stay in the hospital (***Information Week*** 2013). As indicated in this study, technology is expected to continue evolving, with more and more ways for consumers to monitor their health and the health of their loved ones. With the evolution of the technology, healthcare stakeholders will look toward health informaticist to provide them with the guidance to incorporate appropriate health technology into daily use.

Personal Health Records

A PHR is "an electronic or paper health record maintained and updated by an individual for himself or herself; a tool that individuals can use to collect, track, and share past and current information about their health or the health of someone in their care" (AHIMA 2017). PHRs have been promoted as one method to improve health for patients and reduce the costs of healthcare (Hillestad et al. 2005). The Meaningful Use Stage 2 requirements which include portals to the EHR data were able to drive up the development of a tethered-PHR (CMS 2016b). This requirement was widely implemented in 2018 to fulfill one objective from the Meaningful Use Stage 3 requirements, in which to allow patient to use electronic means to access their health data for the purpose of viewing, downloading, and transmitting their PHI. Timely access and the capability of the patient to use an application of their choice from the provider's certified software were required to meet this objective (CMS 2016a). Because of this effort, PHRs became an important consumer tool to enable patients' ability to manage their own health and health information and increase patient engagement.

page 416

The PHR can be a paper record or health information stored on a smart card, USB drive, CD, or computer software such as Apple HealthKit. Apple HealthKit collects data from health and fitness apps like Fitbit and provides easy access to information via a dashboard of data. It also allows access to others at the individual's discretion (Apple 2016). The electronic PHR (ePHR) can be a stand-alone record or tethered to an EHR maintained by a health facility, such as a hospital or health insurance organization or a system such as the US Department of Veterans Affairs. The PHR is not in common usage, which brings up the question of what the barriers are to its adoption. A systematic literature study conducted in 2015 identified the most commonly stated barriers to PHR implementation to be:

- Risk of security and confidentiality and potential for breaches.
- Physicians question the quality, accuracy, and integrity of patient-reported data
- Costs and who would bear the costs. Costs to the providers include those for the technology and licensing fees, time for checking records, and increased workload to maintain.
- Health literacy of individuals seeking information and concern for their inability to completely understand the information.
- Computer literacy and difficulties for individuals to navigate systems (Vance et al. 2015)

A study of patients who participated in the pilot of My HealtheVet, the PHR system of the US Department of Veterans Affairs, examined the patients' experience reading their health information online (Woods et al. 2013). The overall results of the study were positive, with three themes emerging to reflect what they perceived to be the benefits of having this access. The first benefit that respondents perceived is the positive effect on their communication with providers and the health system. The second benefit is that respondents agreed that their knowledge of their health can be improved with the information from PHR, which allowed them to improve their self-care as well. The third benefit perceived by respondents was that the PHR facilitated increased participation in their own care by allowing them to follow-up on abnormal test results and allowing them to have their voice heard in decision-making of their own care. The same study also identified issues to the use of the PHR. The first issue revolved around the initial stress at seeing their own health information, some related to discomfort with the language that was used (due to the specific medical terms being used) or perceived errors or

inconsistencies in the information (between what they feel or experienced compared to what was being written in the PHR). The second issue highlighted the difficulty in using the technology to access the PHR. This issue was expressed by a few participants, mostly during the pilot phase of the program (which were addressed in the final version of the PHR). The third issue focused on the discomfort from respondents that preferred to not knowing the level of information provided within the PHR. These respondents noted that they had not seen or heard this type of information before. Overall, the results were predominantly positive for patients to have and use this access to their PHI, but there was a perceived need for providers to use easily understandable medical language to provide information for their patients.

Disparities in access to health information when it is tied to computer technology is also an issue for the increased use of PHRs and web portals. Patients in underserved areas or those without high-speed Internet access, a computer at home, or a smartphone might find themselves unable to take advantage of the latest health information technology advances. A study in Northern California of a group of more than 14,000 diabetes patients, with demographic characteristics similar to the overall population in Northern California, found that older patients were less likely to use the site and found a pervasive racial and educational disparity in use of the computer portal. The study expressed a fear that "those most at risk for poor diabetes outcomes may fall further behind as health systems increasingly rely on the Internet and limit current modes of access and communications" (Sarakar et al. 2011). Based on this finding, we can make a conclusion that although it is good to develop new uses of technology and IT to help patients, it is important for health information professionals to consider how those who do not possess the latest technology can also improve and maintain their health using information from more traditional means.

Some of the benefits of using electronic PHRs relate to the interoperability of having an electronic PHR tethered to the electronic medical record. Project Health Design is a national program, funded by grants from foundations, designed to stimulate the development of new tools for personal health management by using the PHR (Brennan et al. 2007). The plan is to create a set of prototype applications to support interoperable resources for managing health challenges. If successful, the future of the PHR is bright and vibrant. PHRs have been noted to be "fragmented, non-scalable, without data standards, shared terminologies or common architectures" (Brennan et al. 2007). The usability in the design of applications was felt to be important for full realization of the benefits for healthcare and making the PHRs more usable in the management of chronic diseases.

The idea that the use of PHRs will play a key role in enabling the electronic health environment is becoming more common. The PHR has the potential to create a more balanced and complete view of the patient, especially when patients have multiple physicians or are unable to care for themselves. The patient is referred to as a consumer, rather than as a patient, to reflect this mindset change and to underscore the fact that health records are no longer the sole property of the physician or hospital.

The verdict on the uses and usability of PHRs is far from final; however, especially for patients who suffer from a chronic disease or have multiple providers, they can prove to be very helpful. AHIMA and the AHIMA Foundation maintain the MyPHR.com website, providing an overview of PHRs, as well as guidance on how to initiate a PHR for consumers and their loved ones (AHIMA 2016b). As consumers become increasingly tech-savvy and as the population continues to age, understanding and managing personal health information will become ever more important.

There have continued to be advances in technology, new regulatory requirements for patient access to their information, and an increased consumer need for and interest in

more and complete information so that they can be more in charge of their information. Consumers now have access through portals to test results and other health information, and they have another way to communicate with their providers. New means of control of their own health are available through mobile devices and applications (Butler 2015). Related to this is the concept of patient-generated health data (PGHD), which are "health-related data created, recorded, or gathered by or from patients (or family members or other caregivers) to help address a health concern" (HealthIT 2016). PGHD differs from data compiled by practitioners in the clinical setting in that the patients are responsible for capturing the data and they also are the ones to determine how it is to be shared or used.

Patients provide their health information to the healthcare professionals as a part of seeking treatment whether it was recorded or not. How the information is captured, transmitted, and stored, however, is evolving. The PGHD consists of health history or similar information recorded by the patient by way of a portal, the biometric data generated from home health monitoring equipment, or assorted kinds of lifestyle information, such as data obtained from fitness devices or recorded on mobile device apps such as those used for weight management or diabetes monitoring (AHIMA Workgroup 2015). Support for incorporating the PHGD into the patient's EHR seems to be increasing with the continued focus on patient engagement and value-based care, a plan whereby payment for care is related to the quality and efficiency of care (AHIMA 2017) and its potential use in Meaningful Use. Including PGHD into the EHR has legal implications, specifically when considering the defined legal health record, risks of using the information, and all of the record management and information management considerations (AHIMA Workgroup 2015).

Smooth Integration into Daily Living: Internet-of-Things and SmartCare Concept

CHI is evolving at an exceedingly fast pace due to the rapid pace of technology innovations. Advancement in computer sciences and broadband networks have fueled a revolution in medical device innovation. Multiple stand-alone devices can easily be integrated into multifunction mobile devices, allowing more portability and easier access to tools needed for daily living. Miniaturized healthcare via handheld and wearable technology can seamlessly integrate into the data environment using IoT. Robotics technology enables haptic feedback, facilitating physical contact over distance.

IoT refers to a network of physical objects, loaded with sensing technology, software, brainware, and other communication technology for the purpose of allowing data exchange with other devices and systems that are connected to the object via network, such as the internet. IoT devices range from sophisticated objects powered by artificial intelligence (such as Alexa, Google Home, and so on) to common household objects (such as smart refrigerators, smart TVs, smart lights, and more). Experts estimate that 19 billion such devices will be part of the IoT by 2025 (Statista 2022).

The interconnectivity of these common devices around the home allows digital health to be delivered seamlessly into a person's environment. One of the scenarios of the use of IoT for digital health is captured in the SmartCare concept. In SmartCare, the home is equipped with IoT devices to facilitate independent living, allowing patient and caregiver support, driven by cognitive artificial intelligence to understand the context of the people living in the home. Virtual care can be used to directly connect the home to the clinic on demand using common household media devices, such as smart TVs or voice assistant devices, with sensors around the home providing information for clinicians to triage and determine the best course of action for the health of the individual at home. mHealth can be combined with the IoT at home to complete the health wellness ecosystem, allowing individuals to continuously monitor their own health and adjust or personalize their care

to fit with their lifestyle. Data can be aggregated at health command centers, where healthcare providers can monitor their patients' health needs and direct care as needed just-in-time. The aggregated data would also allow retail healthcare to better manage their service and inventory, allowing time-efficient and cost-efficient care.

Check Your Understanding 14.2

Answer the following questions.

1. What are the two forms of interactivity that can be found in digital health?
   a. Synchronous and asynchronous
   b. Direct and indirect
   c. Personal and community
   d. Physical and virtual
   e. Symmetrical and asymmetrical
2. What has been reported as a challenge in telemedicine?
   a. Successful communication over the distance
   b. Feelings of depersonalization
   c. Convenient access to specialized services
   d. Improved professional education
   e. Reduction in health care costs, for example, from readmissions
3. What would be considered as key mobile health capabilities?
   a. Capability to transmit large images such as x-rays and MRI scans
   b. Connectivity with health information exchange
   c. Reliable communication channel that will always be available
   d. Context sensitive, allowing personalization of intervention to fit with patients' situations and conditions
   e. Providing sensitive health information to peer network via social networking

4. Which are not commonly used technology for CHI?
   a. Bulletin boards and online discussion groups
   b. Wearable sensors such as glucose monitors
   c. Medical devices
   d. Smartphone apps
   e. Internet of Things
5. What would be a correct statement about CHI?
   a. CHI focuses on patient and consumer engagement with the goal of empowering consumers to take ownership of their health information, which could promote healthier lifestyles.
   b. CHI facilitates deeper understanding between patient and provider using technology such as mobile apps and smartphones.
   c. CHI allows providers and healthcare entities stronger control over patient information because the platform facilitates exchange or "push" of information from provider and healthcare entity.
   d. CHI connects patient with any available third-party vendor through the internet, and therefore cannot be relied on as a source of health information
6. Which of these statements is true of PHRs?
   a. Efficacy of PHR does not depend on a patient's level of education and computer literacy.
   b. Healthcare providers can be contacted at any time to provide explanations of the information stored within.

c. Individuals can manage their own health information to provide when needed or when occasions arise (such as in an emergency situation).
d. PHRs are always connected with an EHR system used by providers.
e. Health information can be easily exchanged through a health information exchange organization.

7. The SmartCare concept aims to do what?
   a. Facilitate independent living by enhancing patient and caregiver support with cognitive technology supported by artificial intelligence
   b. Bridge IT companies to healthcare
   c. Permit a more intrusive but medical grade care or services at home
   d. Allow healthcare providers to provide support via robotic or sensing technology
   e. Connect individuals with peers and community support at home.
8. In regard to telehealth's interactivity, which statement fits into what is currently known about IT?
   a. Due to the advancement of technology, synchronous interactivity in telemedicine is inexpensive and requires simple information and communication technology.
   b. Asynchronous interactivity can replace synchronous interactivity due to the fast speed of the bandwidth.
   c. Synchronous activity is performed in real time but only allows very limited types of communication, specifically only video and audio communication.
   d. Asynchronous interactivity can be implemented using low-cost technology.
   e. Both synchronous and asynchronous interactivity are available only on higher-end telemedicine or telehealth solutions.

Provide a short answer for the following.

9. Who published the ***Clinician's Guide to Use of Telehealth for CRT Service Provision*** in 2021?
10. What technology can be used as part of mHealth?
11. What is the term for the ability to assess behavior when it is happening in real time?

Challenges in Health and Digital Health Solutions

It is important to know what **health information** health stakeholders are exploring to understand the challenges in health as well as developing the digital health solutions to address them. The Pew Survey group began researching consumer behaviors related to online health behaviors in the early 2000s. During 15 years of study, they noted similarities regarding consumers' online searches and the health information they sought. A comprehensive 2014 study reported in the ***Journal of Medical Internet Research*** compiled data for a number of variables related to searches and search topics (Lee et al. 2014). A 2012 study conducted by Pew revealed that the information most frequently accessed by consumers was related to a specific disease or medical problem (55 percent), medical treatment or procedures (43 percent), weight control or loss (27 percent), and health insurance (25 percent; Pew Internet and American Life Project 2013). These study findings validate the need for a variety of kinds of timely and easily accessible information to promote the patients in being active participants in their own or a family member's healthcare. Utilization of the data and information also need to be measured to understand the challenge in the information pathway as well as the impact and benefit of the information being delivered, a topic that has been little investigated so far.

It is essential that the healthcare stakeholders know the availability and the reasonable life span of the information, as well as when to update or retract their information. For example, consumers are seeking online health information about time-sensitive topics such as recalls. This time-sensitive information requires the availability of dynamically updated online factual content. In contrast, complex or chronic health conditions do not require minute-to-minute updates, but incremental changes as new information becomes available.

Quality of information plays an important role in building trusts from health stakeholders in digital health as well as ensuring that care and service can be delivered and consumed appropriately. Lack of quality information could lead to incorrect decision-making, which could create a poor patient experience. Fractured information might cause insufficient continuity of care and inadequate supportive acts (due to lack of understanding on patient context and condition). Instructions that are difficult to understand could reduce patients' and providers' adherence to guidelines for care and intervention. Information also needs to be presented appropriately to stakeholders in regard to their understanding of their health context. Acceptability of information depends on whether the information aligns with the users' norms, individual beliefs, and practices. Personalization of the message to deliver health information by fitting the message into this cognitive structure would allow health information to be consumed more easily.

Understanding the utilization of information allows health stakeholders to improve the efficiency of the service to deliver care. Outcome measurement processes allow identification of which services and solutions are used the most, at what time, and for what purpose. In addition, outcome measurement processes will also allow health stakeholders to understand solutions with low adherence or low utilization, and the reasons for these difficulties. Geographic information allows an understanding of locations that need the most service but have access limitations. Ultimately, improving the efficiency of healthcare will have a positive impact on the cost of delivering care as manual processes are streamlined, resources are allocated more effectively, and payer mechanisms can be coordinated with evidence to justify the services rendered.

Digital health can also address the issue of accountability. Audit trails and activity logs could give insight on why some patients are not adhering to certain interventions or are not showing sufficient engagement with their providers. Feedback from digital health solutions could help stakeholders to become more aware of their behavior and act to make improvements accordingly. Transparency of access logs to both patients and providers could show how data are being viewed and used, improving security and ensuring that PHI is being accessed appropriately.

Roadmap for the Future

As mentioned earlier in the chapter, the road toward digital health adoption depends on the development of safe, effective, and widely accessible digital health technology, which depends on methods used to accumulate knowledge of the digital technology and its use in healthcare. Outcome measurements allow the generation of useful data and development of evidence in support of digital health. The results would also help identify and build clear causal models that address what healthcare scenario would be appropriate to deliver through digital health as well as what cannot be delivered due to the limitations of the technology employed in digital health. Aggregation of outcomes could also help identify the essential active components of digital health technology that make intervention work. This evidence would then influence the policy for reimbursement of services, leading to sustainable revenue models for digital health.

Several themes that have been identified as barriers to the adoption of digital health by healthcare entities include issues related to technology, financial, legal, business strategy, and human factors. LeRouge and Garfield presented methods that can be used to address these barriers, including identifying organizational champions for driving business strategy, enabling data sharing and integration with other health entities, increasing the awareness of digital health in the community, developing realistic return on investment projections and formulas for revenue, cost savings (such as the savings from avoiding penalties for not preventing or providing appropriate service), and exploring expansion of services while following and advocating digital health-related regulations (LeRouge and Garfield, 2013).

Every health stakeholder can promote digital equity. An article published in ***JAMA*** (Rodriguez et al. 2020) outlines a roadmap that elaborates on how each health stakeholder, no matter what their role, can improve technology access, digital health literacy, and inclusive design. Patients can advocate for their preferences in digital health. Providers can create and use innovative digital health tools and protocols, collect outcome measures, and improve awareness of digital health through advanced training and education. Administrators and policymakers can continue the push to make permanent changes in reimbursement policies to promote equity and access while also facilitating permanent digital health licensure and interstate licensing for providers. Data scientists and researchers need to involve patients and caregivers in the design of their new digital health innovations, promote the use of uniform data sets and outcome measures, carry out studies in digital health with more rigorous designs to develop evidence of digital health, and use implementation science approaches to measure the dose and exposure, access, cost, and quality of care impacts of digital health.

Chapter 14 Review Exercises

Answer the following questions.

1. Which of the following organizations identified the four stakeholders in digital health?
   a. Office of the National Coordinator for Health Information Technology
   b. Agency for Healthcare Research and Quality
   c. Centers for Disease Control and Prevention
   d. World Health Organization
2. Addressing this challenge of digital health would build up trusts of health stakeholders in digital health as well as ensuring that care and service can be delivered appropriately.
   a. Availability of digital health
   b. Quality of digital health information
   c. Accountability of digital health
   d. Utilization of digital health
3. What would not be considered a strategy to address barriers of adoption of digital health in a healthcare entity?
   a. Enabling data sharing and integration with other health entities
   b. Developing realistic return on investment projections and formulas for digital health revenue
   c. Mitigating risk and cost savings by avoiding implementation of digital health
   d. Exploring expansion of services while following and advocating digital health-related regulations

4. What types of health information are not available to consumers via a search on the internet?
    a. Medicine or medical device information
    b. Medical conditions
    c. Lifestyle tips
    d. Public health event history reports

Discussion

5. What are the expectations of patients as stakeholders of digital health?
6. What are the main differences between telemedicine and mHealth?
7. Provide the reasoning why mobile technology is appropriate and can deliver new types of intervention for healthcare.
8. What is a personal health record?
9. What is the difference between a personal health record and a patient portal?
10. How can we promote digital equity?

REFERENCES

Agency for Healthcare Research and Quality (AHRQ). 2022. Delivering Health Care Quality Information Through Mass Media. https://www.ahrq.gov/talkingquality/distribute/media/mass-media.html. Retrieved on November, 2022

Al-Alusi MA, Ding E, McManus DD, Lubitz SA. 2019. Wearing Your Heart on Your Sleeve: the Future of Cardiac Rhythm Monitoring. Curr Cardiol Rep. 2019 Nov 25;21(12):158. doi: 10.1007/s11886-019-1223-8. PMID: 31768764; PMCID: PMC7777824.

American Health Information Management Association (AHIMA) Consumer. 2017. Pocket Glossary for Health Information Management and Technology. 5th ed. AHIMA Press.

American Health Information Management Association (AHIMA) Workgroup. 2015. Including patient-generated health data in electronic health records. Journal of AHIMA 86(2):54–57.

American Health Information Management Association (AHIMA). 2016a. Consumer Engagement Toolkit. Chicago, IL: AHIMA; 2016

American Health Information Management Association (AHIMA). 2016b. "myPHR." https://www.myphr.com/.

American Medical Association (AMA,). 2018. "Digital health implementation playbook series." https://www.ama-assn.org/practice-management/digital/digital-health-implementation-playbook-series. (accessed on July 4, 2022).

American Medical Association (AMA,). 2021. Telehealth Implementation Playbook. American Medical Association. Available at: https://www.ama-assn.org/practice-management/digital/telehealth-implementation-playbook-overview.

American Telemedicine Association (ATA,). 2018. "American Telemedicine Association Core Operational guidelines for telehealth services involving provider-patient interactions." https://www.americantelemed.org/resources/core-operational-guidelines-for-telehealth-services-involving-provider-patient-interactions (accessed on July 2, 2022).

American Telemedicine Association (ATA,). 2022. "American Telemedicine Association Policy Council Federal and State Policy Activities Tracker." https://www.americantelemed.org/policy/policy-council/.

Apple. 2016. "Health: An innovative new way to use your health and fitness information." http://www.apple.com/ios/health/.

Barbosa, W., K. Zhou, E. Waddell, T. Myers, and E. R. Dorsey. 2021. Improving access to care: Telemedicine across medical domains. Annual Review of Public Health, 42, :463–481.

Bashshur, R. L., E. A. Krupinski, J. H. Thrall, and N. Bashshur. 2016. The Empirical Foundations of Teleradiology and Related Applications: A Review of the Evidence. Telemedicine J and E -Health. 2016 Nov; 22(11):868–898. doi:10.1089/tmj.2016.0149. Epub 2016 Sep 1. PMID: 27585301; PMCID: PMC5107673.

Brall, C., P. Schröder-Bäck, and E. Maeckelberghe. 2019. Ethical aspects of digital health from a justice point of view. European Journal of Public Health, 29 (Supplement_. 3), :18–22.

Brennan, P. F., S. Downs, and D. Kenron. 2007. Project health design: Stimulating the next generation of personal health records. In AMIA Annual Symposium Proceedings 2007. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2655909/.

Breton, M., Sullivan, E.E., Deville-Stoetzel, N. et al. 2021. Telehealth Challenges During COVID-19 As Reported By Primary Healthcare Physicians in Quebec and Massachusetts. BMC Fam Pract 22, 192 (2021). https://doi.org/10.1186/s12875-021-01543-4

Butler, M. 2015. Mastering the inbox information era: Patient-Generated Data and Mobile Health Are Changing the Management of Health Information. Journal of AHIMA 86(9):18–21.

Centers for Medicare and Medicaid Services (CMS). 2016a. "EHR Incentive Programs in 2015 through 2017 Patient Electronic Access." https://www.cms.gov/.

Centers for Medicare and Medicaid Services (CMS). 2016b. "Electronic Health Records (EHR) Incentive Program." https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/index.html?redirect=/ehrincentiveprograms .

Clinician Task Force. 2021. Clinician's Guide to Use of Telehealth for CRT Service Provision. https://cliniciantaskforce.us/assets/Telehealth%20CRT%20Service%20Provision%20Guidelines%20January%202021.pdf . Retrieved on July, 2022

Crabb, Stephanie. 2019. "Digital Health: Promise, Perils, and Perspective." Journal of AHIMA 90, no. 5 (May 2019): 32-35.

Dorsey, E. R., and E. J. Topol. 2020. Telemedicine 2020 and the next decade. The Lancet, 395(10227), :859.

Eysenbach, G. 2001. What is e-health? Journal of Medical Internet Research, 3(2), :e833.

Fact, MR (FMR). 2022. "U.S. Electronic Health Records Market." https://www.factmr.com/report/us-electronic-health-records-market. (accessed on July 10, 2022).

Food and Drug Administration (FDA). 2019. "Device Software Functions Including Mobile Medical Applications." Available at: https://www.fda.gov/medical-devices/digital-health-center-excellence/device-software-functions-including-mobile-medical-applications (accessed on January 20, 2022). Retrieved on July, 2022

Forbes Business Insight. 2021. Telemedicine Market Reports. https://www.fortunebusinessinsights.com/industry-reports/telemedicine-market-101067. Retrieved on October, 2022

Gajarawala, S. N., and J. N. Pelkowski. 2021. Telehealth Benefits and Barriers. J Nurse Pract. 2021 Feb;17(2):218-221. doi: 10.1016/j.nurpra.2020.09.014. Epub 2020 Oct 21. PMID: 33106751; PMCID: PMC7577680.

Gunasekeran, D. V., R. M. W. W. Tseng, Y. C. Tham, and T. Y. Wong. 2021. Applications of Digital Health for Public Health Responses to COVID-19: A Systematic Scoping Review of Artificial Intelligence, Telehealth and Related Technologies. NPJ Digital Medicine, 4(1), :1–6.

Guo, C., H. Ashrafian, S. Ghafur, G. Fontana, C. Gardner, and M. Prime. 2020. Challenges for the evaluation of digital health solutions—A call for innovative evidence generation approaches. NPJ Digital Medicine, 3(1), :1–14.

Hall, E. T. 1963. "A System for the Notation of Proxemic Behavior". American Anthropologist. 65 (5): 1003–1026. doi: 10.1525/aa.1963.65.5.02a00020

Healthcaredive, 2020. 'Weeks where decades happen': Telehealth 6 months into COVID-19. https://www.healthcaredive.com/news/telehealth-6-months-coronavirus/581447/. Retrieved on October, 2022

HealthIT.gov. 2016b. "Patient-Generated Health Data." https://www.healthit.gov/policy-researchers-implementers/patient-generated-health-data.

Hillestad, R., J. Bigelow, A. Bower, F. Girosi, R. Meili, R. Scoville, and R. Taylor. 2005. Can electronic medical record systems transform healthcare? Potential health benefits, savings, and costs. Health Affairs 24(5):1103–1117.

Information Week. 2014. "Fitbit gadget could improve hospital care." http://www.informationweek.com/mobile/fitbit-gadget-could-improve-hospital-care/d/d-id/1111478.

Iyawa, G. E., M. Herselman, and A. Botha. 2016. Digital health innovation ecosystems: From systematic literature review to conceptual framework. Procedia Computer Science, 100, :244–252.

Jetty, A., Y. Jabbarpour, M. Westfall, D. B. Kamerow, S. Petterson, and J. M. Westfall. 2021. Capacity of primary care to deliver telehealth in the United States. The Journal of the American Board of Family Medicine, 34(Supplement.), S48–S54.

Kostkova, P. 2015. Grand Challenges in Digital Health. Frontiers in Public Health, 3, :134.

Kukafka, R. 2019. Digital Health Consumers on the Road to the Future. Journal of Medical Internet Research, 21(11):e16359.

Lee, H., A. Kreshnik, J. Hughes, and L. Emmerton. 2014. Dr Google and the consumer: A qualitative study exploring the navigational needs and online health information-seeking behaviors. Journal of Medical Internet Research: 16(12):e262.

Lehne, M., J. Sass, A. Essenwanger, J. Schepers, and S. Thun. 2019. Why digital medicine depends on interoperability. NPJ Digital Medicine, 2(1), :1–5.

LeRouge, C., and M. J. Garfield. 2013. Crossing the telemedicine chasm: Have the U.S. barriers to widespread adoption of telemedicine been significantly reduced? International Journal of Environmental Research and Public hHealth, 10(12),):6472–6484. https://doi.org/10.3390/ijerph10126472.

Loumbas, E. Z., and M. Peters. 2022. Data Security Implications in Digital Health: A Cautionary Tale. In Nursing and Informatics for the 21st Century–Embracing a

Digital World, 3rd Edition., Book 4, (pp. 109–128. New York: Productivity Press.
Mobisoft. 2018. E-Health Is Going To Change Healthcare. https://mobisoftinfotech.com/resources/blog/e-health-is-going-to-change-healthcare/. Retrieved on July, 2022
Naeem, I., H. Quan, S. Singh, N. Chowdhury, M. Chowdhury, V. Saini, and T. Tc. 2022. Factors Associated With Willingness to Share Health Information: Rapid Review. JMIR Hum Factors. 2022 Feb 9;9(1):e20702. doi: 10.2196/20702. PMID: 35138263; PMCID: PMC8867291.
Neprash, H. T., A. Everhart, D. McAlpine, L. B. Smith, B. Sheridan, and D. A. Cross. 2021. Measuring Primary Care Exam Length Using Electronic Health Record Data. Medical Care: January 2021 - Volume 59 - Issue 1 - p 62-66 doi: 10.1097/MLR.0000000000001450
Norman, C. D., and H. A. Skinner. 2006. eHealth Literacy: Essential skills for consumer health in a networked world. Journal of Medical Internet Research 8(2):e9. doi: 10.2196/jmir.8.2. e9.
Nuzzi, R., G. Boscia, P. Marolo, and F. Ricardi. 2021. The Impact of Artificial Intelligence and Deep Learning in Eye Diseases: A Review. Front Med (Lausanne). 2021 Aug 30;8:710329. doi: 10.3389/fmed.2021.710329. PMID: 34527682; PMCID: PMC8437147.
PC Magazine. 2015. 10 Apps That Are Changing Healthcare. https://www.pcmag.com/news/10-apps-that-are-changing-healthcare. Retrieved on July, 2022
Pew Internet and American Life Project. 2014. "Health Online." http://www.pewinternet.org/files/old-media/Files/Reports/PIP_HealthOnline.pdf.

Rodriguez, J. A., C. R. Clark, and D. W. Bates. 2020. Digital Health Equity as a Necessity in the 21st Century Cures Act era. JAMA. 2020;323(23):2381–2382. doi:10.1001/jama.2020.7858.
Sarakar, U., A. Karter, J. Liu, N. Adler, R. Nguyen, A. Lopez, and D. Schillinger. 2011. Social disparities in internet patient portal use in diabetes: Evidence that the digital divide extends beyond access. JAMIA 18(3):318–321.
Statista. 2022. IoT Connected Devices Worldwide. https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/. Retrieved on October, 2022
Stokel-Walker, C. 2020. Why telemedicine is here to stay. BMJ, 371:m3603.
United States Department of Health and Human Services (HHS). 2021. New HHS Study Shows 63-Fold Increase in Medicare Telehealth Utilization During the Pandemic. https://www.hhs.gov/about/news/2021/12/03/new-hhs-study-shows-63-fold-increase-in-medicare-telehealth-utilization-during-pandemic.html . Retrieved on October, 2022
van Panhuis, W. G., P. Paul, C. Emerson, J. Grefenstette, R. Wilder, A. J. Herbst, D. Heymann, and D. S. Burke. 2014. A systematic review of barriers to data sharing in public health. BMC Public Health. 2014 Nov 05;14(1):1144. doi: 10.1186/1471-2458-14-1144. https://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-14-1144.1471-2458-14-1144
Vance, B., B. Tomblin, J. Studeny, and A. Coustasse. 2015, March. Benefits and barriers for adoption of personal health records. Marshall Digital Scholar. Spring(3). http://mds.marshall.edu/cgi/viewcontent.cgi?article=1135&context=mgmt_faculty.
Varkey, B. 2021. Principles of Clinical Ethics and Their Application to Practice. Med Princ Pract 2021;30:17-28. doi: 10.1159/000509119

Wahsheh, H. A., and M. S. Al-Zahrani. 2021, April. Lightweight Cryptographic and Artificial Intelligence Models for Anti-smishing. In International Conference on Emerging Technologies and Intelligent Systems, (pp. 483-496). Springer, Cham, Switzerland: Springer.

Woods, S., E. Schwartz, A. Tuepker, N. Press, K. Nazi, C. Turvey, and W. Nichol. 2014. Patient experiences with full electronic access to health records and clinical notes through the My HealtheVet personal health record pilot: Qualitative study. Journal of Medical Internet Research 15(3):e65. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3636169/.

World Health Organization (WHO). 2019a. "WHO guideline: Recommendations on Digital Interventions for Health System Strengthening." World Health Organization. http://apps.who.int/iris/bitstream/handle/10665/311941/9789241550505-eng.pdf.

World Health Organization (WHO). 2019b. "Classification of digital health interventions." World Health Organization. https://www.who.int/publications/i/item/WHO-RHR-18.06.

ADDITIONAL RESOURCES

Aichner, T., and F. Jacob. 2015. Measuring the degree of corporate social media. ***International Journal of Market Research 57***(2):257–275.

Centers for Medicare and Medicaid Services (CMS). 2015. "CMS fact sheet: EHR incentive programs in 2015 and beyond." https://www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2015-Fact-sheets-items/2015-10-06-2.html .

HealthIT.gov. 2016. "***What is a patient portal?"*** https://www.healthit.gov/providers-professionals/faqs/what-patient-portal.

Kimball, A. B., K. C. Corey and J. C. Kvedar. 2015. Engaging patients to decrease costs and improve outcomes. ***Medical Economics 92(2):36–37.*** http://medicaleconomics.modernmedicine.com/medical-economics/news/engaging-patients-decrease-costs-and-improve-outcomes .

Nazi, K. L. 2010. "***VA's Blue Button: Empowering people with their data"*** [Web log]. http://www.blogs.va.gov/VAntage/866/vas-blue-button-empowering-people-with-their-data/.

Patil, D. J., and M. Uppaluru. 2015. "Celebrating the 5 year anniversary of Blue Button and open health data." https://www.whitehouse.gov/blog/2015/10/01/celebrating-5-year-anniversary-blue-button-open-health-data .

The Precision Medicine Initiative. 2016. https://www.whitehouse.gov/precision-medicine.

US Department of Veterans Affairs. 2014. "***New VA Blue Buttons features for 2013:*** VA Notes, and CCD and more." http://www.va.gov/bluebutton/.

Chapter

15

Public Health Informatics

By Lesley Clack, ScD, CPH

Learning Objectives

- Explain the importance of public health informatics
- Distinguish between public health informatics and population health informatics
- Identify the three core functions of public health
- Understand the 10 essential services of public health
- Explain the competencies needed for public health informatics professionals

KEY TERMS

Assessment
Assurance
Globalization
Health outcomes
Infodemic
Infodemic management
Policy development
Population health
Population health informatics
Public health
Public health informatics
Public Health Information Network
Public health surveillance

Informatics is a discipline that is applicable to many different fields in healthcare. Biomedical and health informatics (HI) have developed their own areas of emphasis for their professions and disciplines (AMIA 2022). Public health informatics is an area of practice within the field of health informatics. The Centers for Disease Control and Prevention (CDC) defines public health as "the science of protecting and improving the health of people and their communities" (CDC 2022). The CDC defines public health informatics (PHI) as "the systematic application of knowledge about systems that capture, manage, analyze and use information to improve population-level health outcomes" (CDC 2021). Health outcomes are the specific indicators used to measure health status. The field of public health informatics focuses on the application of informatics in public health areas, such as surveillance, prevention, preparedness, and health promotion (AMIA 2022).

Core Areas and Essential Services for Public Health

Public health informatics is a systematic application of information science, computer science, and technology in the practice of public health that serves three core functions of public health: the assessment of population health, policy development, and assurance of the availability of high-quality public health services (Yasnoff et al 2000). The three core functions of public health are embedded within the 10 essential services for public health (Nelson and Staggers 2018). The core functions and essential services for public health are shown in figure 15.1.

Figure 15.1. Three core functions and ten essential services for public health

Core function	Essential services
Assessment	Monitor health status to identify community health problems
	Diagnose and investigate health problems and health hazards in the community
Policy development	Inform, educate, and empower people about health issues
	Mobilize community partnerships to identify and solve health problems
	Develop policies and plans that support individual and community health efforts
Assurance	Enforce laws and regulations that protect health and ensure safety
	Link people to needed personal health services and assure the provision of healthcare when otherwise unavailable
	Assure a competent public health and personal healthcare workforce
	Evaluate effectiveness, accessibility, and quality of personal and population-based health services
	Research for new insights and innovative solutions to health problems

Source: IOM (1988)

Public health informatics is different from informatics in other fields in that it involves "a focus on applications of information science and technology that promote the health of populations, rather than of individuals; a focus on disease prevention, rather than treatment; a focus on preventive intervention at all vulnerable points in the causal chains

leading to disease, injury, or disability; and operation typically within a governmental, rather than a private, context" (Magnuson and Fu 2016, 4). The Office of Disease Prevention and Health Promotion (2022) has set an objective to enhance the use and capabilities of informatics in public health. The objective is listed as a high-priority public health issue.

Public Health Informatics Competencies

The competencies for Public Health Informaticians, shown in figure 15.2, were created in 2009 through a joint effort between the CDC and the Association of Schools and Programs of Public Health (Public Health Foundation 2021). These competencies were designed to supplement the eight domains within the Core Competencies for Public Health Professionals (shown in figure 15.3) to specifically define competencies for those with roles and expertise in public health informatics. The role of the informatician in public health is to do the following:

- Plan, design, and define functional requirements for public health information systems
- Evaluate the application and impact of information systems in support of health goals
- Serve as a liaison between multidisciplinary teams
- Use data standards to support interoperability of data between systems
- Ensure confidentiality, security, and integrity standards
- Be knowledgeable of health data standards, sources, and meaningful use of health data. (CDC 2009)

The competencies for public health informaticians found in figure 15.2 are intended to provide information regarding the knowledge, skills, and abilities needed for public health informaticians. These competencies care intended to be used for hiring, training, and evaluating public health informaticians. For example, these competencies can be used in job descriptions and as goals in job performance evaluations. The eight domains within the core competencies shown in figure 15.3 reflect the skills areas within public health informatics. For example, data analytics and assessment skills in domain 1 are needed in public health informatics to be able to collect, analyze, and use the vast amount of public health data that is available from a wide variety of sources.

Figure 15.2. Competencies for public health informaticians

Public Health Informatician	Senior Public Health Informatician
Supports development of strategic direction for public health informatics within the enterprise	Leads creation of strategic direction for public health informatics
Participates in development of knowledge management tools for the enterprise	Leads knowledge management for enterprise
Uses informatics standards	Ensures use of informatics standards
Ensures that knowledge, information, and data needs of project or program users or stakeholders are met	Ensures that knowledge, information, and data needs of users and stakeholders are met
Supports information system development, procurement, and implementation that meet public health program needs	Ensures that information systems development, procurement, and implementation meet public health program needs
Manages IT operations related to project or program (for public health agencies with internal IT operations)	Ensures IT operations are managed effectively to support public health programs (for public health agencies with internal IT operations)
Monitors IT operations managed by external organizations	Ensures adequacy of IT operations managed by external organizations
Communicates with cross-disciplinary leaders and team members	Communicates with elected officials, policymakers, agency staff, and the public
Evaluates information systems and applications	Ensures evaluation of information systems and applications
Participates in applied public health informatics research for new insights and innovative solutions to health problems	Conducts applied public health informatics research for new insights and innovative solutions to health problems
Contributes to development of public health information systems that are interoperable with other relevant information systems	Ensures that public health information systems are interoperable with other relevant information systems
Supports use of informatics to integrate clinical health, environmental risk, and population health	Uses informatics to integrate clinical health, environmental risk, and population health
Implements solutions that ensure confidentiality, security, and integrity while maximizing availability of information for public health	Develops solutions that ensure confidentiality, security, and integrity while maximizing availability of information for public health
Conducts education and training in public health informatics	Contributes to progress in the field of public health informatics

Source: HHS (2019)

Figure 15.3. Core competencies for public health professionals

Domain 1: Data Analytics and Assessment Skills
1.1 Describes factors that affect the health of a community
1.2 Accesses existing quantitative and qualitative data
1.3 Collects quantitative and qualitative data
1.4 Analyzes quantitative and qualitative data
1.5 Manages quantitative and qualitative data
1.6 Uses quantitative and qualitative data
1.7 Applies public health informatics in using data, information, and knowledge
1.8 Assesses community health status
Domain 2: Policy Development and Program Planning Skills
2.1 Develops policies, programs, and services
2.2 Implements policies, programs, and services
2.3 Evaluates policies, programs, services, and organizational performance
2.4 Improves policies, programs, services, and organizational performance
2.5 Influences policies, programs, and services external to the organization
2.6 Engages in organizational strategic planning
2.7 Engages in community health improvement planning
Domain 3: Communication Skills
3.1 Determines communication strategies
3.2 Communicates with internal and external audiences
3.3 Responds to information, misinformation, and disinformation
3.4 Facilitates communication among individuals, groups, and organizations
Domain 4: Health Equity Skills
4.1 Applies principles of ethics, diversity, equity, inclusion, and justice
4.2 Engages in continuous self-reflection about one's biases
4.3 Recognizes the diversity of individuals and populations
4.4 Reduces systemic and structural barriers that perpetuate health inequities
4.5 Implements organizational policies, programs, and services to achieve health equity and social and environmental justice
4.6 Contributes to achieving and sustaining a diverse, inclusive, and competent public health workforce
4.7 Advocates for health equity and social and environmental justice
Domain 5: Community Partnership Skills
5.1 Describes conditions, systems, and policies affecting community health and resilience
5.2 Establishes relationships to improve community health and resilience
5.3 Maintains relationships that improve community health and resilience
5.4 Collaborates with community members and organizations
5.5 Shares power and ownership with community members and others

Domain 6: Public Health Sciences Skills
6.1 Describes systems, policies, and events impacting public health
6.2 Applies public health sciences in delivering the 10 Essential Public Health Services
6.3 Uses evidence in developing, implementing, evaluating, and improving policies, programs, and services
6.4 Contributes to the evidence base for improving health
Domain 7: Management and Finance Skills
7.1 Describes factors that affect the health of an organization
7.2 Secures human resources
7.3 Manages human resources
7.4 Engages in professional development
7.5 Secures financial resources
7.6 Manages financial resources
7.7 Implements organizational policies, programs, and services to achieve diversity, equity, inclusion, and justice
7.8 Manages programs and services
7.9 Engages in contingency planning
7.10 Applies critical thinking in decision-making
7.11 Engages individuals and teams to achieve program and organizational goals
7.12 Facilitates collaboration among individuals, groups, and organizations
7.13 Engages in performance management
Domain 8: Leadership and Systems-Thinking Skills
8.1 Creates opportunities to achieve cross-sector alignment
8.2 Implements a vision for a healthy community
8.3 Addresses facilitators and barriers impacting delivery of 10 Essential Public Health Services
8.4 Creates opportunities for creativity and innovation
8.5 Responds to emerging needs
8.6 Manages organizational change
8.7 Engages politicians, policymakers, and the public to support public health infrastructure
8.8 Advocates for public health

Source: Public Health Foundation (2021)

Check Your Understanding 15.1

Answer the following questions.

1. The assessment core function of public health includes which of the following essential services?
   a. Diagnose and investigate health problems and health hazards in the community
   b. Develop policies and plans that support individual and community health efforts
   c. Research for new insights and innovative solutions to health problems

d. All of the above
e. None of the above

2. The assurance core function of public health includes all of the following essential services except: page 438
   a. Evaluate effectiveness, accessibility, and quality of personal and population-based health services
   b. Link people to needed personal health services and assure the provision of healthcare when otherwise unavailable
   c. Mobilize community partnerships to identify and solve health problems
   d. Enforce laws and regulations that protect health and ensure safety
3. The policy development core function of public health includes all of the following essential services except:
   a. Assure a competent public health and personal healthcare workforce
   b. Develop policies and plans that support individual and community health efforts
   c. Inform, educate, and empower people about health issues
   d. Mobilize community partnerships to identify and solve health problems
4. Which of the following is not a competency for a senior public health informatician?
   a. Ensures adequacy of IT operations managed by external organizations
   b. Participates in development of knowledge management tools for the enterprise
   c. Conducts applied public health informatics research for new insights and innovative solutions to health problems
   d. Leads knowledge management for enterprise
5. Which of the following is not one of the core competencies for public health under the health equity domain?
   a. Recognizes the diversity of individuals and populations
   b. Applies principles of ethics, diversity, equity, inclusion, and justice
   c. Engages in continuous self-reflection about one's biases
   d. Creates opportunities for creativity and innovation

Case Study: Public Health Infrastructure

Healthy People 2030 was developed by the Office of Disease Prevention and Health Promotion at the US Department of Health and Human Services as a science-based framework that provides 10-year national objectives to improve health and well-being over the next decade. Healthy People 2030 includes 355 core objectives in addition to developmental and research objectives. Public health infrastructure is one of the domains included in Healthy People 2030. The goal of this domain is to make sure that all public health agencies have the appropriate infrastructure for key public health services. This infrastructure includes up-to-date data and information systems that provide the necessary data for responding to public health needs. One of the specific objectives of this domain is to increase the proportion of immunization information systems that track adult immunizations across the life span. This objective is classified as developmental, meaning that it is a high priority for public health and there are evidence-based interventions to address it that do not yet have enough data for analysis. Another objective is to enhance the use and capabilities of informatics in public health, including data sharing, data exchange, and page 439

application to practice, and to use the information for decision-making. This objective is classified as research, meaning that it is a high priority for public health but there are not yet evidence-based interventions available to address it (HHS 2022a).

The Role of Public Health Data

The collection and accessibility of public health data are essential for public health informatics. Public health data may come from a variety of sources, such as mortality statistics, notifiable diseases reporting, laboratory data, ongoing population surveys, and environmental exposure data (Hedberg and Maher 2018). The National Center for HIV, Viral Hepatitis, STD, and TB Prevention (NCHHSTP) Program Collaboration and Service Integration (PCSI) adapted 10 guiding principles for data collection, storage, sharing, and use to ensure security and confidentiality of public health data (2014; Lee and Gostin 2009).

1. Public health data should be acquired, used, disclosed, and stored for legitimate public health purposes.
2. Programs should collect the minimum amount of personally identifiable information necessary to conduct public health activities.
3. Programs should have strong policies to protect the privacy and security of personally identifiable data.
4. Data collection and use policies should reflect respect for the rights of individuals and community groups and minimize undue burden.
5. Programs should have policies and procedures to ensure the quality of any data they collect or use.
6. Programs have the obligation to use and disseminate summary data to relevant stakeholders in a timely manner.
7. Programs should share data for legitimate public health purposes and may establish data-use agreements to facilitate sharing data in a timely manner.
8. Public health data should be maintained in a secure environment and transmitted through secure methods.
9. Minimize the number of persons and entities granted access to identifiable data.
10. Program officials should be active, responsible stewards of public health data. (NCHHSTP PCSI 2014; Lee and Gostin 2009)

Public Health Information Network

The Public Health Information Network (PHIN) is an initiative developed by the CDC for establishing and implementing a framework for the electronic sharing of public health information. The PHIN was created to establish a standardized way for governments to exchange public health data securely.

Creating a public health information system requires the following five steps:

1. Vision and System Planning. Objective: Envisioning solutions, opportunities, and application of information technology in public health.
2. Health Data Standards and Integration. Objectives: Defining and designing health data standards and transformation (for example, HL7, ICD, SNOMED) and health

domain integration (for example, EHR, CMS, HIE, surveillance, demographics, and social media); and defining and implementing databases tables, data formats, and keys for linking tables and data to support defined health data standards and integration.

3. Data Privacy and Security. Objectives: Defining and implementing health data privacy and HIPAA regulations; and implementing and enforcing data, systems, and communication security.
4. Systems Design and Implementation. Objectives: Defining and designing methods for public health functions, data elements, data flow, case definitions, and message mapping; and implementing information technology for defined functions, data elements, data flow, and case definitions.
5. Visualization, Analysis, and Reporting of Health Data. Objectives: Visualization and implementation of the required analysis, reporting, and meaningful use of the data collected and managed by the system (CDC 2014).

The CDC maintains several value sets in a system called the Public Health Information Network's Vocabulary Access and Distribution System (PHIN VADS; CDC 2021). The PHIN VADS is a web-based vocabulary system that provides a standard vocabulary to the CDC and its public health partners in one place. There are currently 215 code systems, 1,944 value sets, 103 views, and more than 5 million concepts in the PHIN VADS (CDC 2021). An example of some of the frequently used PHIN VADS content is shown in figure 15.4.

Figure 15.4. Frequently used PHIN VADS content

Topic	Description
RCMT	Reportable Condition Mapping Table (RCMT) provides mapping between conditions and its associated lab test and results
ELR HL7 2.5.1	Value sets associated with Electronic Laboratory Reporting (ELR) HL7 2.5.1 Implementation Guide
Immunization HL7 2.5.1	Value sets associated with Immunization HL7 2.5.1 Implementation Guide
Syndromic Surveillance for Emergency Department (ED) and Urgent Care (UC)	Value sets associated with Syndromic Surveillance PHIN messaging guide
Reporting to Central Cancer Registries	Value sets associated with Implementation Guide for Healthcare Provider Reporting to Central Cancer Registries (Release 1.0)
NHSN Healthcare Associated Infection (HAI)	All the value sets associated with National Healthcare Safety Network (NHSN) Healthcare Associated Infection (HAI) Release 5
PHLIP Flu Lab Report Vocabulary	All the value sets associated with PHLIP Flu lab report HL7 message implementation guide

Source: CDC (2021)

Public Health Surveillance

As discussed previously, the Institute of Medicine (IOM) established assessment as one of the core functions of public health and recommended that every public health agency regularly collect and analyze information regarding the health of the community (Thacker et al. 2012). This is a type of public health surveillance. Public health surveillance has been defined as "the systematic, ongoing collection, management, analysis, and interpretation of data followed by the dissemination of these data to public health programs to stimulate public health action" (Thacker et al. 2012, 3). The CDC and the Agency for Toxic Substances and Disease Registry mapped out their Vision for Public Health Surveillance in the 21st Century (Buehler 2012). The vision outlined the following health information needs for the twenty-first century:

- Preparedness
- New information systems and technologies
- Availability of more and new data, new data sources, and databases
- Healthcare reform
- Electronic health and medical records
- Performance accountability
- Comparative effectiveness of medical interventions
- Measuring positive indicators
- Infrastructure and capacity
- Globalization
- Interoperability
- Changes in the economy
- Expanding scope of public health practice
- Data privacy issues
- Data security
- Public participation in public health activities (Hall et al. 2012)

To address these needs, two basic principles were identified: (1) the purpose of public health surveillance, which serves to address a defined public health population or question and to use the data to guide efforts that will protect and promote population health; and (2) the nature of public health surveillance as an ongoing and systematic implementation of a set of processes consisting of planning a system design, data collection, data analysis, interpretation of results of analysis, dissemination and communication of information, and application of information to public health programs and practice (Hall et al. 2012).

The National Notifiable Disease Surveillance System (NNDSS) is a nationwide collaboration in which all local, state, territorial, federal, and international agencies of public health share health information for monitoring, control, and prevention of the occurrence and spread of nationally notifiable infectious diseases (HHS 2022b). The data are collected weekly and evaluated annually by the Council of State and Territorial Epidemiologists and the CDC (HHS 2022b).

Case Study: Real-World Application of Public Health Informatics

An example of public health informatics in action is CDC's FluView. CDC's FluView was a clear-cut way to share national influenza data. These data were needed by the public health community, clinicians, scientists, and the general public. CDC's FluView was a weekly influenza surveillance report prepared by the influenza division (CDC 2013).

In 2020, the CDC created a similar application for the tracking of the coronavirus. The COVID Data Tracker provides daily updates on COVID-19 community levels to determine the impact and actions needed for communities (CDC 2022). Many states also created their own dashboards to track state-level data related to COVID-19. One example is the Georgia Department of Public Health COVID-19 dashboard, which provides data by county for the state of Georgia (Georgia Department of Public Health 2022).

Check Your Understanding 15.2

Answer the following questions.

1. A flu outbreak has occurred in 10 states across the country. To increase knowledge of the health threat, the CDC uses computer science, technology, and applied information that will inform the nation's population. This is an example of:
   a. Public health informatics
   b. Surveillance
   c. Privacy and security
   d. Research
2. Public health informatics uses public health knowledge to:
   a. Broaden the public health knowledge base through learning
   b. Improve population health in daily practice
   c. Further knowledge in public health research
   d. All of the above
3. In reference to creating a public health information system, which of the following does a public health informatics professional consider first when identifying technologies to use for sharing national tuberculosis data?
   a. Health data standards and integration
   b. Vision and systems planning
   c. System design and implementation
4. Public health informatics is used to create programs such as CDC's FluView and COVID-19 Data Tracker. Which of the following three disciplines must work together to visually represent the data in an effective method?
   a. Technology, computer science, and applied information methods
   b. Computer science, epidemiology, and public health
   c. Technology, surveillance systems, and epidemiology
5. Which of the following is not a function of a public health informatician?
   a. Uses data standards to support interoperability of data between systems
   b. Designs, implements, and administers database architecture, privacy, security, and backup procedures
   c. Ensures confidentiality, security, and integrity standards
   d. Is knowledgeable about health data standards, sources, and meaningful use of health data

Population Health Informatics

While public health informatics and population health informatics are both areas of informatics focused on improving the health of the public, they are distinctly different. Where public health serves to promote and protect the health of people and their community, population health "describes the health needs of a defined group's entire life span" and "is a linking thread to help understand the determinants of health of populations" (Joshi et al. 2019, 5). Relative to public health, population health is less associated with public health departments, is more than the sum of its individual parts, includes a broader array of the social determinants of health, and has outcomes that go beyond government public health agencies and health systems (Joshi et al. 2019). The intersection of public health informatics and population health informatics (Gamache et al. 2018) is shown in figure 15.5.

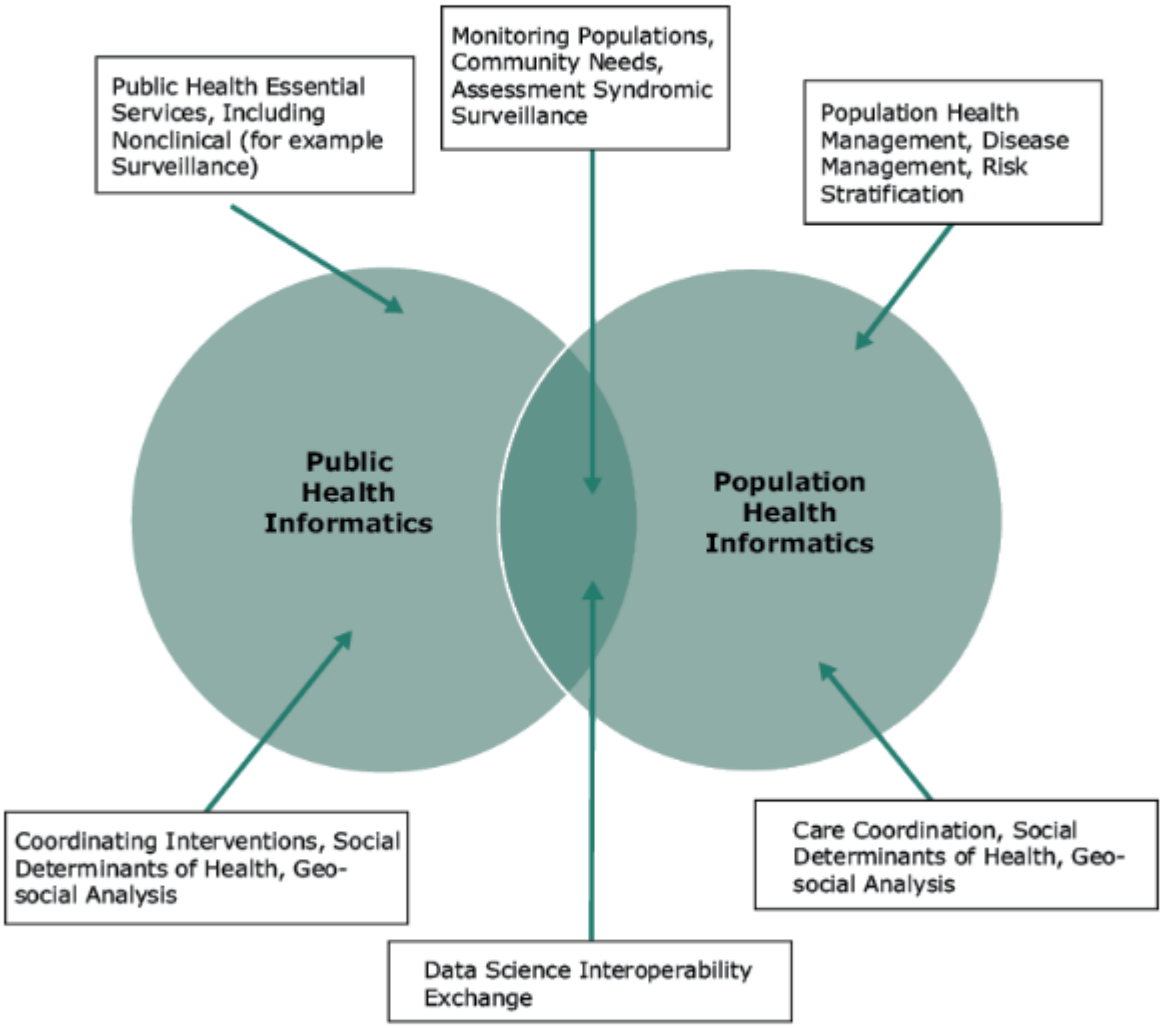

Public Health Essential Services, Including Nonclinical (for example Surveillance)
Monitoring Populations, Community Needs, Assessment Syndromic Surveillance
Population Health Management, Disease Management, Risk Stratification
Public Health Informatics
Population Health Informatics
Coordinating Interventions, Social Determinants of Health, Geo-social Analysis
Data Science Interoperability Exchange
Care Coordination, Social Determinants of Health, Geo-social Analysis

Figure 15.5. Intersection between public health informatics and population health informatics

▶ Expand/collapse Extended Description

Source: Gamache et al. 2018

Globalization of Public Health Informatics

In the wake of the COVID-19 pandemic, the World Health Organization (WHO) published a public health research agenda for managing "infodemics" shown in figure 15.6 (WHO 2021). The purpose of this agenda is that it creates a framework for public health professionals to use globally. Public health informatics has entered into a globalization, meaning that it is now operating on an international scale. An infodemic has been defined as "too much information including false or misleading information in digital and physical environments during a disease outbreak" (WHO 2022, 1). Infodemics lead to confusion and risky behaviors that can affect individuals' health, lead to mistrust in public authorities, and undermine the public health response. Infodemic management refers to the systematic use of risk-based and evidence-based analysis of data to manage the infodemic and reduce its impact on health behaviors during pandemics and epidemics. The goals of infodemic management are to listen to community concerns and questions, promote understanding of risk and health expert advice, build resilience to misinformation, and engage and empower communities to take positive actions (WHO 2022).

page 445

page 446

Figure 15.6. Public health research agenda for infodemic management

Stream 1: Measuring and monitoring the impact of infodemics during health emergencies
1.1 Standardize taxonomies and classifications
1.2 Develop new metrics to measure and quantify infodemics
1.3 Analyze and triangulate data from multiple sources
1.4 Improve evaluation approaches for infodemic interventions
Stream 2: Detecting and understanding the spread and impact of infodemics
2.1 Understand how information originates, evolves and spreads on different platforms and channels
2.2 Assess the role of actors, influencers, platforms, and channels
2.3 Understand how misinformation affects behavior in different populations
2.4 Develop regulatory and ethical principles to mitigate the spread and propagation of harmful health information
Stream 3: Responding and deploying interventions that protect against the infodemic and mitigate its harmful effects
3.1 Design a behavioral change model applicable to infodemic management
3.2 Intervention design for different levels of action to mitigate the infodemics.
Stream 4: Evaluating infodemic interventions and strengthening resilience of individuals and communities to infodemics
4.1 Develop interventions that address individual, community, cultural, and societal factors affecting trust and resilience to misinformation
4.2 Understand and learn from the way misinformation has affected behavior among different populations and in different contexts for specific infodemics
4.3 Identify factors associated with successful infodemic management by health authorities, the media, civil society, the private sector, and other stakeholders
Stream 5: Promoting the development, adaptation, and application of tools for managing infodemics
5.1 Use implementation research evidence in program improvement and policy development
5.2 Promote evidence-based interventions and approaches between countries
5.3 Improve effectiveness and response times to the infodemic during acute health events

Source: WHO (2021)

Chapter 15 Review Exercises

1. Explain why public health informatics is an important area of practice within health informatics.
2. Of the three core public health functions, which one do you think is the most important? Explain your reasoning.
3. Of the 10 essential services for public health, which one do you think is the most important? Explain your reasoning.
4. Compare and contract public health informatics and population health informatics.
5. Why is the PHIN VADS important? What could happen if this type of system did not exist?

6. If you were hiring for a public health informatics professional, which of the competencies would be most important to you and why?
7. Explain what an infodemic is and why that is a relevant concern today.
8. Of the adapted 10guiding principles for data collection, storage, sharing, and use, which do you think are most important and why?
9. Explain the importance of public health surveillance.
10. Explain the importance of the Public Health Information Network.

REFERENCES

American Medical Informatics Association (AMIA). 2022. "Informatics: Research and practice." https://amia.org/about-amia/why-informatics/informatics-research-and-practice.

Buehler, J. W. 2012. CDC's vision for public health surveillance in the 21st century: Introduction. ***MMWR Supplement*** 61(1):1–2.

Centers for Disease Control and Prevention (CDC). 2009. ***Public Health Informatics Competencies.*** Atlanta, GA: US Department of Health and Human Services, CDC. http://www.cdc.gov/informaticscompetencies/.

Centers for Disease Control and Prevention (CDC). 2013. ***FluView.*** Atlanta, GA: US Department of Health and Human Services, CDC. http://gis.cdc.gov/grasp/fluview/main.html.

Centers for Disease Control and Prevention (CDC). 2015. Introduction to Public Health. In ***Public Health 101 Series.*** Atlanta, GA: U.S. Department of Health and Human Services, CDC. https://www.cdc.gov/training/publichealth101/informatics.html.

Centers for Disease Control and Prevention (CDC). 2021. "Public Health Informatics Fellowship Program (PHIFP)." https://www.cdc.gov/phifp/overview/index.html#:~:text=Public%20health%20informatics%20is%20the,program%20in%20public%20health%20informatics .

Centers for Disease Control and Prevention (CDC). 2022. "COVID data tracker." https://covid.cdc.gov/covid-data-tracker/#datatracker-home.

Department of Health and Human Services (HHS). 2009. "Competencies for public health informaticians." https://docplayer.net/3408501-U-s-department-of-health-and-human-services-centers-for-disease-control-and-prevention-office-of-workforce-and-career-development.html .

Department of Health and Human Services (HHS). 2022a. "Healthy People 2030: Public health infrastructure." https://health.gov/healthypeople/objectives-and-data/browse-objectives/public-health-infrastructure .

Department of Health and Human Services (HHS). 2022b. "National Notifiable Diseases Surveillance System (NNDSS)." https://health.gov/healthypeople/objectives-and-data/data-sources-and-methods/data-sources/national-notifiable-diseases-surveillance-system-nndss#:~:text=The%20National%20Notifiable%20Disease%20Surveillance%20System%20%28NNDSS%29%20is,notifiable%20infectious%20and%20some%20noninfectious%20diseases%20and%20conditions .

Gamache, R., H. Kharrazi, and J. P. Weiner. 2018. Public and population health informatics: The bridging of big data to benefit communities. ***Yearbook of Medical***

Informatics 27(1):199–206.
Georgia Department of Public Health. 2022. "Georgia DPH COVID-19 status report." https://ga-covid19.ondemand.sas.com/.
Hall, H. I., A. Correa, P. W. Yoon, and C. R. Braden. 2012. Lexicon, definitions, and conceptual framework for public health surveillance. ***MMWR Supplement*** 61(1):10–15.
Hedberg, K., and J. Maher. 2018. Collecting Data. In ***The CDC Field Epidemiology Manual.*** https://www.cdc.gov/eis/field-epi-manual/chapters/collecting-data.html.
Institute of Medicine (IOM). 1988. ***The Future of Public Health.*** Washington, DC: National Academy Press.
Joshi, A., L. Thorpe, and L. Waldron. 2019. ***Population Health Informatics: Driving Evidence-Based Solutions into Practice***. Burlington, MA: Jones & Bartlett Learning.
Lee, L. M., and L. O. Gostin. 2009. Ethical collection, storage, and use of public health data: A proposal for national privacy protection. ***JAMA*** 302:82–84.
Magnuson, J. A., and P. C. Fu. 2016. ***Public Health Informatics and Information Systems,*** 2nd ed. London: Springer International.
National Center for HIV, Viral Hepatitis, STD, and TB Prevention (NCHHSTP) Program Collaboration and Service Integration (PCSI). 2014. "Ten guiding principles for data collection, storage, sharing and use to ensure security and confidentiality". https://www.cdc.gov/nchhstp/programintegration/tenguidingprinciples.htm
Nelson, R., and N. Staggers. 2018. ***Health Informatics: An Interprofessional Approach.*** St. Louis, MO: Elsevier.
Office of Disease Prevention and Health Promotion. 2022. "Enhance the use and capabilities of informatics in public health- PHI-R06." https://health.gov/healthypeople/objecives-and-data/browse-objectives/public-health-infrastructure/enhance-use-and-capabilities-informatics-public-health-phi-r06 .
Public Health Foundation. 2021. "Core competencies for public health professionals." http://www.phf.org/resourcestools/Documents/Core_Competencies_for_Public_Health_Professionals_2021October.pdf .
Thacker, S. B., J. R. Qualters, and L. M. Lee. 2012. Public health surveillance in the United States: Evolution and challenges. ***MMWR Supplement*** 61(1):3–9.
World Health Organization. 2021. WHO public health research agenda for managing infodemics. Geneva, Switzerland: WHO.
World Health Organization. 2022. Infodemic. Infodemic (who.int)
Yasnoff, W. A., P. W. O'Carroll, D. Koo, R. W. Linkins, and E. M. Kilbourne. 2000. Public health informatics: Improving and transforming public health in the information age. ***Journal of Public Health Management and Practice*** 6(6):67–75.

ADDITIONAL RESOURCES

Centers for Disease Control and Prevention (CDC). 2013. "Environmental exposures." https://www.cdc.gov/nchs/fastats/environmental-exposures.htm.
Centers for Disease Control and Prevention (CDC). 2018. "Introduction to public health informatics." https://www.cdc.gov/training/publichealth101/informatics.html.

Centers for Disease Control and Prevention (CDC). 2022. "National Notifiable Diseases Surveillance System (NNDSS)." https://www.cdc.gov/nndss/index.html.

Centers for Disease Control and Prevention (CDC). 2022. "National public health performance standards." https://www.cdc.gov/publichealthgateway/nphps/index.html.

Centers for Disease Control and Prevention (CDC). 2022. "Public health surveillance and data." https://cdc.gov/surveillance/index.html.

Kharrazi, H., E. C. Lasser, W. A. Yasnoff, J. Loonsk, A. Advani, H. P. Lehmann, D. C. Chin, and J. P. Weiner. 2016. A proposed national research and development agenda for population health informatics: Summary recommendations from a national expert workshop. ***Journal of the American Medical Informatics Association*** 24(1):2–12.

Mastrian, K. G., and D. McGonigle. 2017. ***Informatics for Health Professionals.*** Burlington, MA: Jones & Bartlett Learning.

Morawski, T. S. 2022. "Health Informatics." HIMSS. https://www.himss.org/resources/health-informatics.

National Center for Health Statistics. 2022. "Mortality statistics." https://www.cdc.gov/nchs/nvss/deaths.htm.

National Committee for Quality Assurance. 2022. "Population health program accreditation standards." https://www.ncqa.org/programs/health-plans/population-health-program-accreditation/benefits-support/standards/ .

Public Health Informatics Institute. 2022. "Defining public health informatics." https://phii.org/how-we-do-it/defining-public-health-informatics/.

Sahay, S., T. Sundararaman, and J. Braa. 2017. ***Public Health Informatics***. Oxford, UK: Oxford University Press.

USA Facts. 2022. "Population & Demographics." https://www.usafacts.org/data/topics/people-society/population-and-demographics/?utm_source=bing&utm_campaign=ND-StatsData&msclkid=d99f3520adf614c0684e57d33b4a0849 .

World Health Organization. 2017. "Standards or public health information services in activated health clusters and other humanitarian health coordination mechanisms." https://www.who.int/docs/default-source/documents/publications/standards-for-public-health-information-services.pdf?sfvrsn=328fc80e_1 .

Chapter

16

Trends and Emerging Technologies

By Andi Saptono, PhD and Augie Turano, PhD

Learning Objectives

- Identify the latest emerging technology that can practically affect public health and healthcare
- Summarize the implementation of emerging technology in health and healthcare through case scenarios
- Articulate the role of health informatics in the implementation of emerging technology for health and healthcare

KEY TERMS

Accountable care organization (ACO)
Archival system
Artificial intelligence (AI)
Backup
Care coordination
Clinical decision support
Cloud computing
Data consolidation
Data visualizations
Disease management
Evidence-based medicine (EBM)
Meaningful Use
Patient-centered medical home (PCMH)
Point-of-care testing (POCT)
Precision medicine
Predictive modeling

Quality improvement (QI)
Regional Extension Centers (RECs)
Seamless care

As mentioned throughout this book, health informatics involves utilizing communication and information technology along with health information to affect change, which could be measured in quality of care or increased use and understanding of health data. A typical health informatics scenario involves managing a patient's record using EHR technology in conjunction with data captured from wearable or personal devices in addition to what is received from external entities. This combination creates an enhanced and comprehensive health data platform for supporting and delivering care, reporting, and analysis. The increased usage of health informatics is driving the development of new health technology, including technology to assist statistical analysis of large amounts of data as well as data presentation and visualization software. Much has been accomplished with the increased utilization of EHRs but the full potential has yet to be recognized. This chapter introduces a few of the current and future trends in technology for data management, continuing and preventative care, education, employment, and advanced use cases of health informatics such as in using artificial intelligence (AI) software in medical devices.

Current and Emerging Trends and Terms

Many examples currently show how informatics can be utilized to improve patient outcomes and facilitate efficient delivery of healthcare services. Informatics provides expanded provider access to a patient's past health information and more immediate reporting of lab and other diagnostic test results critical to patient diagnosis and treatment. It impacts the computerized physician order entry (CPOE) systems by providing information directly to the physician for decision making when prescribing. This reduces the time needed to determine which medications to order and this results in the patient getting medications sooner with fewer potential errors (for example, wrong medication, dosage, or frequency). Clinical data checks for drug-to-drug interaction can be enhanced further with drug to food, and interaction with allergies. Enhanced computer speed and performance allows more extensive searches for data interactions. Informatics also increases patient engagement through patient portals that provide access to the patient's health record and additional materials. Topics presented throughout this text reflect on the current status and future possibilities of various aspects of health informatics. These include EHR advances (chapter 2), the information exchange (chapter 3), expanded use of databases (chapter 5), benefits to research and data analysis (chapter 7), and expanded uses of health information in data science (chapter 8).

Although the benefits and advances in health informatics are vast and well understood, there are still areas of concern. Patient privacy and confidentiality pose significant challenges due to the increased use of technology and data exchange. In addition, there are many new uses of these data for decision making, planning, and policy making. Another issue of concern includes the initial cost-effectiveness of developing and implementing systems. However, the benefits of fully functional electronic health information systems include improvements in the quality and convenience of patient care, improved accuracy of diagnosis and health outcomes, better care coordination, increased patient participation, and increases in efficiencies and cost savings long term (HealthIT 2015). Questions remain regarding whether the cost benefit should be a consideration

when there are such significant positive benefits for patient care. With the increased use of artificial intelligence (AI) algorithms for disease prediction and diagnosis, there needs to be more transparency with the type of algorithms used, sample sizes, and any biases that exist in the data. Consumerism will continue to be a factor as patients and providers become accustomed to having access to more information. It is becoming more of an expectation that decisions be based on appropriate data to promote optimal results in the provision and planning of healthcare and in the strategic planning and policy making for the future.

Redefining the Health Informatics Domain

Information is used universally throughout healthcare. Specific domains for consideration in formal educational programs are covered in the section on education of this chapter. What seems to be paramount to the identity of the health informatics educational domain is establishing the knowledge content to adequately define what informatics is and what it does. A comprehension of hardware, software capabilities, report generation and use of various algorithms and computational libraries are of prime importance. Also important is the standardization of commonly accepted elements of data, languages, and vocabularies. A key component of data analytics is the ability to use the information generated to move from raw data to drawing conclusions and making decisions. The functions presented here are not meant to be all-inclusive but rather a sampling of areas that will be impacted by the influence of health informatics.

Emerging Technology

Advances and new uses for technology in healthcare are ongoing and will continue to develop. Newer technology ranges from wearables to digital networks. Here are some examples of recently developed technology.

- Digital diagnostic tools that provide access where it might not otherwise be available. For example, Neurotrack is a computer-based Alzheimer's test, measuring eye movement, which might indicate early-onset Alzheimer's disease.
- Neuro-link, Elon Musk's company, is developing brain chip implants that could aid paralyzed individuals with spinal or brain injuries by giving them the ability to control computerized devices with their minds.
- Ultrafast scans that can capture an image of one beat of the heart. Patients who experience a very fast heart rate are typically denied a scan because of their fast heart rate.
- Massive cloud storage of data in a variety of formats stored in data lakes for health institution enterprise reporting for data tracking, reporting, and analysis. Data lakes are similar to data warehouses but have more capability to handle large amounts of varied data (AWS 2022).

- Wearables, which are a technology worn by the patient, are commonly used for tracking fitness levels. Some patients with Parkinson's disease are utilizing wearable devices to determine characteristics of the disease.
- Patient empowerment with web portals, wearable devices, and more at-home testing, which includes correct and dependable medical information that will be assessable from the patient. This will give clinicians patient data from outside the hospital or office setting, providing patients the opportunity to play a lead role in their own healthcare.

Numerous technologies are on the horizon for implementation and future development. The examples given here illustrate the variety of the potential new technologies and if

known, the status of development of the technology is provided (Mesko 2016a, 2016b).

- Real-time data where checking, monitoring, and recording vital signs or other measures is important for healthcare and to determine when intervention is needed. This would require a sensor that recognizes such things as chewing movements, coughing, smoking, and talking which would then record the data (for example, when too much is eaten or when someone is smoking). These reports can then be reviewed by medical providers in real time. This technology could also be used in individual organs to provide real-time data on how they are functioning. Other examples are blue-tooth glucose monitors that alarm when values fall out of range, or pacemakers or defibrillators that react when monitored values require action.
- Artificial intelligence (AI) is a high-level information technology used in developing machines that imitate human qualities such as learning and reasoning (ORACLE 2022). AI systems can have cognitive abilities that can translate information into knowledge for use in better decision-making. For example, IBM Watson provides assistance to physicians in making decisions but is not a substitute for human involvement. Watson displayed some general limitations and did not do well with general diagnosis. In corrective action, recognizing the limitation that Watson had, IBM focused on "Narrow AI", that is, AI aimed at solving smaller problems. They focused on cancer and drug efficacy and had much better success. Similar applications mine data from medical records to provide better and faster service or predict which medications will work in a given situation. Most of these applications are in the initial phases of deployment and are used in select facilities. (Jie Z et al. 2021)
- Gamifying is "the process of adding games or game-like elements to something (as a task) so as to encourage participation" (2010) (Yu-Kai Chou 2022). This can be a motivation for patients to comply with their treatments and communicate information, and for providers to track certain activities. One example of gamification in healthcare is an application called Mango Health. This application encourages patients to take medications by rewarding them with points each time they take their medication resulting in monetary prizes, such as a gift card or a donation to a charity (CopperMobile 2022).

- Holographic data input where screens and keyboards can be projected on the wall or on a table for ease and accessibility in the clinical setting. The data are stored exclusively in a cloud with only a small projector needed to see the data. Also, common appliances like an X-Box can be used to augment physical therapy by showing proper motion on a projection, then having a patient repeat what was seen. Programs can then comment on physical movement or corrective behavior.
- Multifunctional radiology is projected to have the ability to detect many symptoms, biomarkers, and medical problems at one time and is predicted to include imaging techniques and personalized diagnostics with interventions in real-time.
- In silico clinical trials allow small microchips to be used as models of human cells, organs, or systems for simulation testing rather than the long and expensive traditional clinical trials. A current application of this technology, referred to as human organs-on-chips, uses stem cells to mimic organs using a series of devices (Wyss Institute 2022). Predictions are that the clinical trial process will be changed significantly and completely replace animal testing.
- Genomics and the Human Genome Project substantiated the concept of personalized medicine with tailored therapy and tailored drugs. There are already evidence-based medicine applications in existence such as the use of genetic information to

determine dosages of medication in heart patients resulting in a decreased hospitalization rate. With the increased emphasis on personalized medicine, there will be new opportunities for conducting bedside DNA analysis.

- Digestible sensors that function like a thin e-skin to transmit information. These sensors measure and transmit a variety of health parameters such and blood biomarkers, neurological symptoms, and vital signs to the cloud. Alerts will be sent when a medical event occurs and can even call an ambulance when needed.
- Real-time diagnostics and Wearable Devices - Personal health monitoring in real-time has changed significantly in the last five years. Devices like the iPhone and wireless technologies like Bluetooth enable monitoring of many bodily functions continuously. Heart and glucose monitors, Fitbits, and similar devices can track and alert both patients and clinicians that an event may be occurring. Chronic diseases can be treated with pacemakers, defibrillators, closed-loop pancreas systems, insulin pumps et cetera that can process information as it is received and perform an action.
- MetaVerse can allow a combination of leveraging virtual reality (VR) and augmented reality (AR) technology to operate in virtual spaces. The metaverse's great potential can include improved surgical precision, therapeutic usages such as PTSD treatment, social-distancing accommodations, and psychological aversion therapy where patients can interact with situations that cause them anxiety. All of these uses and more will lead to widespread adoption of the metaverse as a disruptive transformation in healthcare (Swider 2022).

Emerging Applications

There are numerous applications on both desktop and mobile platforms for patients, consumers, physicians, and other providers. These applications benefit consumers and healthcare providers by making medical information and disease treatment accessible to the public. There are also engaging tools that inform the general public about the challenges in the field of public health. For example, CDC provides consumers with a free epidemiology application called Solve the Outbreak. This application casts the user in the role of disease detective, also called an Epidemic Intelligence Service Agent, who examines clues and makes decisions about realistic disease outbreak scenarios (CDC n.d.a). One question for the user to ponder might be, "Do you ask for a quarantine or for more diagnostic tests?" Points are earned for correct decisions as the player advances to harder scenarios. It is beneficial to know that these kinds of applications are possible and understanding the data required to develop them.

For physicians, there are numerous CDC applications available on various diseases with current up-to-date information (CDC n.d.b). For example, CDC Grand Rounds provides topic-specific facts on diseases and conditions, family life, life stages, populations, and injuries, to name a few (CDC n.d.c). CDC Grand Rounds addresses public health issues, key challenges, cutting-edge scientific evidence, potential solutions, and recent developments. These applications and corresponding video and images can help physicians and patients, by putting diagnostic and teaching information at their fingertips.

There is also no better example of the public and government working together than the COVID19 pandemic. CDC and many other government institutions were tracking and recording COVID incidents throughout the United States and around the world. A web display and tracking site was setup by CDC that showed the latest statistics on the geographic disease progression and vaccination rates in various parts of the country (CDC n.d.d). Tracking and contact applications sprung up from private institutions as well as entrepreneurs offering sites where a person could get vaccinated with wait times, mortality rates by country and other helpful information. Disease and geographic

information show where virus hot spots are allowing for medication and supplies to target those areas and to potentially restrict travel in and out of those areas until virus spread was under control. An excellent global Covid pandemic visualization can be viewed from an application created at John Hopkins University School of Medicine which displays pandemic data from across the globe. (JHU 2022).

Virtualization: Expanding Storage and Role of Cloud Computing

Health information originates in different data types to include paper charts, documents, structured and unstructured data from electronically generated EHRs, reports, images, faxes, and written notes all increasing at an exponential rate. Even electronic means of storing digital data run into space limitations and must include mechanisms to maintain the privacy and security of the patient information while meeting the HIPAA requirements. Keeping data available and performance less than one second response time is a challenge. There is no best solution to solving the challenge of maintaining an adequate and secure means of storing health information, but storage has become much more affordable, and techniques for copying and reproducing data have also improved.

Valuable information that is stored electronically must be backed up. A backup is a copy that reproduces the original data files onto a secondary device that can be moved offline where it is safe from tampering. The backup provides long-term retention and preservation of the data. The longer a system is used the more data that is accumulated, but too much data online will slow down application performance. Most patient data must be kept for decades so a solution to retaining data and still having it available is to use a separate archival system. An archival system or database provides a historical copy saved at a particular point in time and can be used later to recover and restore the information in a database. Failures can occur as a result of a hardware failure, software bug, tampering, database corruption, or other failure such as a malicious virus. A Picture Archival and Communication Systems (PACS) is a convenient and economical means of acquiring, storing, retrieving, and displaying images originating from multiple sources. PACS systems store images from many departments such as ophthalmology, dermatology, pathology, nuclear medicine, and other clinical domains. Medical imaging such as X-rays, MRI's, CAT scans and more, generally create very large files and are kept in the PACS for specialized handling and storage. With the threat of destructive malware such as ransomware it is critically important to have a solid backup strategy in the event that an enterprise is attacked, and all the data is encrypted maliciously. Having restorable backups can help get an institution back online and not be helpless in threatening situations with the threat of massive data loss.

With most EHR systems when backups are performed, systems can be rendered temporarily unavailable. In a healthcare environment this can cause problems with patient chart availability, ordering tests, or looking at images. There are new technologies that are available with advanced storage devices. A SAN (Storage Area Network) can have features where a point in time "snap-shot" can be performed in under a minute that is restorable in the event of database corruption or other problems. A "snap-shot" can also be converted to a full clone of the database which takes longer but does not require downtime. A "snap-clone" can become a full backup copy of the data and can be done three times faster or more compared to a traditional backup to disk or tape. Having flexibility and high-speed performance allows hospital IT staff to do more frequent backups. When SANs are used, restoration of data is faster and minimizes inconvenience to hospital staff and patients.

Another means of storing information and running applications, which has become more prevalent in healthcare, is cloud computing. The definition of cloud computing is "a

model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (for example, networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction" (NIST 2011). Organizations are turning to the cloud for a few reasons; less expense for hardware and software; greater accessibility of information from multiple locations and devices, and the ability to meet disaster recovery needs. Data stored in the cloud are secured with data encryption and other security measures which addresses the ongoing issue of information security. There are many cloud vendors that offer both the computing platform as a service as well applications hosted in the cloud. Many EHR vendors now host both their data and applications in the cloud. The cloud offers low upfront capital costs, removes the need for physical on-premise data centers and data can be easily "geolocated," meaning that it can be stored and replicated in multiple data centers with a geographic separation of fifty miles or more. The cloud offers scalability as well, although applications must be programmed to take advantage of cloud features such as adding addition virtual machines for more computing power, or multiple storage devices to achieve maximum scalability and throughput. These features offer high levels of performance as well as protection and resiliency which are desirable for protecting healthcare data and for achieving high data availability of healthcare data.

Data consolidation is a process for bringing together and summarizing large amounts of data that can originate from one or more sources. Due to the many mergers of healthcare entities and the coordination of care across larger populations and geographic areas, the need for data consolidation and aggregation has increased. To an extent, EHRs have the capability to consolidate structured data to facilitate the large-scale aggregation processes needed for healthcare mergers or large accountable care organizations. Going forward, there will be greater focus on consolidating the unstructured data present in EHRs and other data formats. Merging data from different EHRs or data sources can cause terminology and database metadata issues; missing data, and potential semantic loss on certain fields where there are different units of measurements, variations in database schemas (e.g., a 30 characters field has to transform into a 20 characters field), or a terminology transformation that might be required to move a prescription or procedure from one vocabulary to another. Generally, this data is stored in a data warehouse or a more modern data lake cloud environment, where the previous issues can be addressed, and data can be cleaned and homogenized.

Genetic data can hold the key to many diseases and treatments. The sheer amount of genetic data generated creates unique situations and challenges in terms of handling complex genomic data sets. Several methods for maintaining the data are being used to reduce the size of the genome sequences and perform searches and analysis. It should be recognized first that the genome sequence for individuals is similar with approximately 99 percent of the sequence identical for everyone. (Alzu'bi et al. 2014). It is therefore redundant to store thousands of these sequences in a database. A better method for doing this is to record one reference genome and then record differences of the other genomes using this reference or a variation of these two means. It was reported that in one situation, the resulting data set was 750 times smaller when reduced this way. Another approach is to use a two-component distributed information system. In this arrangement, an application server manages primarily the metadata and the other component is a data storage system with web interface links to databases where various genomic data sets are maintained (Alzu'bi et al. 2014). Examples of the databases can be found at the National Center for Biotechnology Information (NCBI) website (NCBI 2016). There are many projects now that take the phenotypic information from an EHR and combine it with Genomic information (CDC 2015). This exercise can help determine what medicines

might have better efficacy, or what treatments might have a better chance for success. As more knowledge is gained from the correlation between the genomic, phenotypic information and biomarkers, medicine is transforming into a more detailed, scientific discipline with higher precision and better informed clinicians who can use this data in deciding treatments and hopefully improving projected outcomes.

Improved Continuity of Care and Increased Efficiency

The concepts of efficiency and continuity of care are not new but the capabilities possible when utilizing advanced technology are immense. Programs that promote online communication like Zoom, MS Teams, and Apple's Facetime promote more interaction among providers and patients. Having shared data in a large accessible repository can also increase the flow of information between your hospital and the extended healthcare ecosystem. A central data store, or shared accessible medical application can help stop duplicative testing, provide better patient data tracking, and eliminate procedures just to name a few issues. All of these issues are better for patients and reduce costs. Increased sharing of data coupled with the new initiatives described below, can lead to improvements that can be transformational in terms of continuity of care and efficiency.

Care Coordination

Another concept conceived to improve healthcare in an efficient, cost-effective manner is Care coordination, which is "the organization of your treatment across several health care providers" (HealthIT 2016). Care coordination involves making sure that all members of a patient's medical team and ecosystem know all the treatments, procedures, desires of the patient, social background, medications, allergies, etcetera to provide safe, high value healthcare to the patient. Many times patients are sent out on referrals to specialists but when they arrive the specialist is unclear about the visit. Having better communication, documentation, patient planning, and access to more information can alleviate that situation and improve care. Medical homes and accountable care organizations (ACOs) are two common ways to coordinate care, which are described in the following sections.

Medical Homes

A patient-centered medical home (PCMH), also referred to as medical home is a "model or a philosophy of primary care that is patient-centered, comprehensive, team-based, coordinated, accessible, and focused on quality and safety" (PCPCC 2016). The PCMH model has been used frequently for those patients with chronic diseases such as renal failure, stroke, heart disease, diabetes, and others. (CDC n.d.e) True patient centric care coordination involves domains that may not fall into a traditional patient plan such as dental history, behavior health, social background, patient education and involvement to achieve high quality healthcare. It provides a model of the organizational and delivery aspects of how primary care should **be provided**.

Five functions identified that are key components to a comprehensive Care Coordination program are listed here (AHRQ 2021):

- Comprehensive care: Provision of physical and mental healthcare, involves a team of providers. Information is used to document the care provided and to serve as a communication tool among providers and caregivers.
- Patient-centered: Relationship based, orientation to whole person.
- Coordinated care: Care delivered across entire healthcare system to include hospitals, specialty care, home care, and community services.
- Accessible services: Easier accessibility for primary care (quicker appointments, more options for appointment times) and 24/7 access to team member either by phone or electronically.

- Quality and safety: Commitment to quality improvement with evidence-based medicine and clinical support tools to guide the decision.

Patient-centered medical home encompasses a partnership between not only the clinicians and the patient but also their family who can be involved in the decision making and participate in their medical care. However, it is the provider team that is ultimately accountable for prevention and treatment of both acute and chronic conditions. The "coordination of care" refers to the whole healthcare system to include specialized care, services of hospitals, home health, available community services, and supporting services. Timely access to care with reasonable wait times and extended hours, continual access electronically or by phone, and the use of IT solutions to enhance good communication are all part of accessible care. The commitment to quality and safety relies on providers and staff who are actively engaged in quality improvement to assure that informed decision making is taking place (PCPCC 2016).

Accountable Care Organizations

A definition of an accountable care organization (ACO) is groups of doctors, hospitals, and other healthcare providers, who come together voluntarily to give coordinated high-quality care to their Medicare patients with the goal of ensuring that patients, especially the chronically ill, get the right care at the right time, while avoiding unnecessary duplication of services and preventing medical errors (CMS n.d.a.). However, there are other organizations that are also ACO's such the Veterans Administration hospitals, Kaiser Permanente, Geisinger Health System and any other healthcare organization that have aligned team-based care with high quality healthcare but are not insurance groups. It is important to distinguish an HMO (Health Management Organization) from an ACO; the difference being HMOs are insurance groups that contract with clinicians, while ACOs consist of clinician groups that contract with insurers.

page 461

ACOs are a payment reform model responsible for not only providing care in instances of illness and injury, but also offering the potential to improve quality and promote health and wellness while eliminating waste and improving efficiencies. To do this, those involved with such an organization must have complete patient health information and must truly understand the care and services they provide. Often this understanding can only be obtained with data collection and analysis supported by coordinated IT initiatives. Interoperability and a seamless flow of information are key issues for the ACO to meet its objectives. One change that has been made based on ACO involvement is identifying a key member of the care team, usually a physician, to establish a trust relationship with the patient and to increase patient engagement in their care. Another example is the concerted effort to establish relationships for care coordination with organizations and individuals outside of the ACO who may provide services to the patient such as home health agencies and organizations for specific diseases such as the American Cancer Society. When ACOs are functioning appropriately, there are increased funds in the Medicare Shared Savings Program due to decreased costs in the provision of care, decreased number of emergency department (ED) visits, and lower hospital readmission rates. These results indicate improved quality of care along with financial savings.

ACOs must demonstrate that they can provide high-quality care at a savings. To facilitate this goal, CMS has a new ACO model called the Next Generation Accountable Care Organization (NGACO). The goal of the NGACO model is to pair ACOs that have programs reflecting best practices that combine provisioning of quality care with cost reductions with ACOs that are seeking to achieve that goal (CMS n.d.a.). The core principles of this model are: to allow beneficiaries to get needed services from the providers of their choice; have a program that rewards quality; rewards improvements that create more efficiency and keeps financial stability. The end goal of this model is to

provide high quality care, improve efficiency and spend healthcare dollars more wisely by cutting duplications and sharing knowledge. Most patients who receive care that is provided by multiple services and physicians will generally have experienced frustration with labs, medications, allergies, history and other clinical indicators not being available at the time of care. The NGACO was designed to address those and other issues while providing better healthcare. There is always danger when assigning the nomenclature of "Next Gen" to anything because invariably there will be new version with improvements that negate "Next Gen." CMS has redesigned the older model and replaced it with the ACO Realizing Equity, Access, and Community Health (REACH) model. This model enhances Medicare in underserved communities and make measurable changes to address health disparities. (CMS n.d.b.).

The Department of Veterans Affairs (VA) is an example of an ACO since it provides care to over nine million patients across the United States. The VA provides inhouse care but also refers many patients out to the community for specialized treatments. Small VA hospitals, particularly those in rural areas, might not have an MRI machine, transplant facilities, or cancer specialists so VA will pay to have those treatments performed outside the VA. Congress appropriates a finite budget, and the VA must work within that budget to provide high quality care to all its patients. With more than 170 hospitals and 2,000 clinics, the VA has the power to buy products such as medicine, medical equipment, beds, supplies etcetera in large quantities many times getting a volume discount. It also has centers of excellence where specialized care for spinal cord injuries, prosthetics, and other combat-related illnesses are treated by experienced clinicians focused on these specific conditions and illnesses (VA 2022). With a finite budget the VA must do its best to maximize healthcare for its patient by instituting as many preventive screenings, patient education and clinical follow-ups as possible, to avoid fully manifested disease states which are the most expensive to treat.

Quality Improvement

Quality improvement (QI) includes the activities that measure the quality of a service or product through systems or process evaluation and then implements revised processes that result in better healthcare outcomes for patients, based on standards of care (AHIMA 2017). Quality improvement processes never stop, a constant iteration should occur until a solid framework with predictable outcomes is achieved. An example of healthcare quality improvement might be to study data from post-operative surgeries to try to minimize infections with new processes or techniques. The lessons learned from QI can often be included in more formal processes sometimes referred to as "process mapping". Health records usually are data obtained from the EHR are a primary source for the QI function to identify problems, services performed and then analyze the outcomes of care that have been provided in the organization. Many IT solutions are evident today in QI processes such as alert devices, new practices of addressing medication dispensing through better use of the EHR, and better methods of communication among providers. An example of this might be an alert sent to a physician that has prescribed Ibuprofen for pain. This might cause harm because the latest lab values show abnormal renal functions which may not have a diagnosis because the disease has not gotten to the critical stage. This kind of QI alert and safety measure is healthy for the patient and can save resources from healthcare entities.

A process that is expected to be more widely used in the future is modeling, which can help data analysts identify patterns in observed data. Unlike prescriptive analytics, which uses data sets to help enhance existing processes and improve operational performance, predictive modelling uses machine learning and artificial intelligence models to discover correlations across disparate data sources and provide actionable recommendations

(Bonderud 2021). Results of any predictive analytics are vastly improved with high quality data. Anomalous or messy data will lead to aberrant results particularly with Machine Learning/AI algorithms. Data can be cleaned after collection but if many clinicians or researchers are using the same data, that process will be repeated needlessly. Imposing early constraints when data is first captured reduces the effort needed for the study of data later. A related area is point-of-care testing (POCT) where diagnostic tests are performed at or near where the patient care is taking place. POCT is done primarily for the convenience and support of the patient. Simple X-rays, quick blood tests, or quick microscopic analyses are great examples. This kind of testing improves the timeliness of the results generation, which could result in higher quality care because a diagnosis can be done quickly while the patient is still present. Quick results can save trips to the lab or more visits to the hospital or other service providers. Similarly, data management systems that eliminate the need for manual data entry or multiple entry of the same information (patient hospital Medical Record number) have the potential to significantly improve the reliability of the data by eliminating human errors. Use of Bluetooth or USB technology for migration of data from medical devices can both speed data acquisition and improve data accuracy and quality. Many instruments and medical devices allow the direct interface of their output to the EHR for viewing by clinicians.

Aggregation of Data to Support Decision-Making for Population Health

The definition of population health has continued to evolve over time. In recent years, a threefold purpose has emerged to include improving individual care, reducing the cost of care, and improving the health of populations. Initiatives supporting population health are increasing due to the movement toward value-based care. Activities to support population health initiatives can require significant investment of both hardware or software architecture and data engineering from a business need perspective. Gathering data from several service providers and healthcare institutions where data models vary as do database field names and metadata, requires data transformation and careful patient matching. Data can be put into a data warehouse or population health platform where data analysis can be performed with a wider spectrum of data for a more comprehensive patient view.

From a health education perspective, the Division of Population Health from the CDC has been charged with managing public programs that provide cross-cutting, chronic disease, and health promotion expertise (CDC n.d.f). This group administers several programs on topics such as arthritis, Alzheimer's disease and healthy aging, excessive alcohol use, epilepsy, lupus, and other disease by creating brochures and educational material about the basics of the disease, resources available and special topics, respectively. There are also food and nutrition, physical activity, and safety applications. Likewise, CMS released educational brochures to patients with chronic illnesses like kidney disease, diabetes, and heart disease. These informative brochures offer treatment guidance, diet, and health relevant information. Reading can help a patient improve their outcomes with self-education. If a patient is aware of social, physical or dietary changes and how it might affect their health, they have the potential to make behavioral changes and achieve better health.

Data gathering and analysis of clinical data requires collection of information from many clinical data sources. EHRs, data from community care, and data gathered from patient portals can offer clinicians and researchers a wider array of data than was ever available before. However, this data may be in a variety of data formats, be semantically encoded using several varying standards, and results can have variations in lab tests, pharmacy terminologies, diagnoses, procedures codes, and any other data where values can use multiple coding systems. Several coding systems are used to capture data

associated with different medical elements such as medications, procedures, diagnosis, lab tests, etcetera. ICD-10-CM and ICD-10-PCS are clinical classification systems that organize data to facilitate retrieval but are not designed for data input or primary documentation. Having a clinician try to code various terminologies at the time of input would not only create errors because of lack of clinical terminology and mapping expertise; but slow down input. SNOMED CT, RxNorm, Multum, and LOINC are clinical terminology systems derived from health informatics. These systems codify clinical information captured during the course of care. Clean and uniform data is a better data source for analytics. Complex systems must have the capability of storing large amounts of data (coded data, metadata, and content). Once data is normalized, cataloged, and tracked, it can be used for clinical, operations or research to build health population reports, measure compliance and be used as source of data for applications.

Preventive Care

The goal of preventive care is to avoid disease by getting regular checkups, vaccinations, and diagnostic screenings such as mammograms and colonoscopies regularly. Full blown disease is more difficult to treat, costs more financially and emotionally, and is harmful to the patient. When disease presents itself, treatment involves all of the interventions that occur in response to a disease or condition. With the onset of disease, there are the consequences to a person's health and well-being in addition to the higher costs associated with treatment as opposed to maintenance of good health. Consistent with the increased capabilities to use information in the provision of care, new initiatives are being introduced that support the focus on preventive care. Several of these are discussed below.

Precision Medicine

Precision medicine is an innovative approach that involves tailoring medical treatments that take into account genetics, diet, environment, and lifestyle. Precision medicine is customized to the individual characteristics of each patient through the ability to identify and classify subpopulations that differ in their susceptibility to a specific treatment (An and Vodovotz 2015). Analysis of big data, such as seeking evidence for an earlier diagnosis of a specific disease, is fundamental to the precision medicine process to develop the tactics to prevent and detect disease and to maintain a healthy life. Millions of people have already been touched by the area of precision medicine that has grown directly from biomedical research. The development and the evolving nature of the COVID vaccines are evidence of precision medicine, where transferring the instructions stored in mRNA to make the proteins to combat the invading virus. Disciplines such as precision oncology are now identifying the molecular fingerprints of various cancers, immunotherapy, Pharmacogenomics, and rare diseases. (NIH 2020). Once the fragments are identified the molecular structure can be used to create blocking molecules to degrade or kill the abnormal cells. One result of using precision medicine is the detection of gene pattern activity that could lead to a blood test that will quickly and accurately detect the presence of sepsis (a deadly, systemic inflammation). Many times, a cut or inflammation if not treated properly can lead to severe infection and could result in the loss of a limb. A recent study on the predictors of death in adults ages 57 to 85 found that the psychosocial well-being, frailty, and similar factors are better predictors of death rather than using standard measures such as having hypertension or diabetes (Okwerekwu 2016). The data and communication resources needed for this type of data analysis are available with the current technology but gathering the data from social sources and family has remained a challenge. Detection of sentiment from Internet posts, or garnering data from patient, friends and family can produce data that is unclear and lead to poor data analysis.

University of California, Davis highlights the four characteristics of precision medicine, comprised in 4Ps of Precision Medicine (UC DAVIS 2022). These characteristics can help us understand the important aspect that should be kept in mind when developing solutions for precision medicine:

- Personalized – Personalized medicine should consider the uniqueness of an individual's personal profile, such as their genetic makeup, environment, etc. Care decisions should integrate known illness factors and the patient's unique characteristics using many data sources to personalize care.
- Proactive or preventive - Precision medicine can "upstream" public health or population health interventions and make them available on an individual level. For example, public health educations to reduce risk of disease should be customized dynamically to an individual's context when delivered via a precision medicine solution to help the individual understand the content and can relate to it better.
- Participatory – In precision medicine, each individual should be allowed to own their health and care. Communities can define important targets of wellness and illness and be full participants in research and clinical improvement.
- Predictive – Through establishing the personalized mapping of all health factors including individual's uniqueness, the most appropriate treatment can be proposed to optimize the response rates and reduce side effects for the individual.

Research on the IT approach and challenges of data integration of clinical and genomic data in precision medicine was conducted with individuals such as Chief medical informatic officers, physicians, and biomedical and pathology directors. Twenty-nine percent responded that they were conducting precision medicine within their facility (HIMSS 2016). Precision medicine is primarily done on-site at large medical centers because of available funding, technology, and expertise. The focus of precision medicine has been on cancer although other conditions and diseases are also studied. The success rate for precision medicine related to cancer is about seven percent so there is much more research to be done (Daily Advisory 2018). The reasons for the high number of facilities focusing on cancer is that it is a genomic disease where learning about and understanding the genetic changes can lead to treatment. Precision medicine funding has been provided primarily to the National Cancer Institute (NCI) for the development of cancer treatment (HIMSS 2016). Targeted therapies based on genomic data can help determine drugs that could be used successfully from a wider array of therapies. Small molecule drugs are tiny enough to enter a cancer cell once they find it and disrupt the cells activity. Advances in drug – cell interactions and determinations of molecular structure help clinicians and researchers focus on individual therapies based on a patient's unique DNA.

Disease Management

Disease management (DM) "emphasizes the provider–patient relationship in the development and execution of the plan of care, prevention strategies using evidence-based guidelines to limit complications and exacerbations, and evaluation based on outcomes that support improved overall health" (AHIMA 2017). Treatment for most diseases have a decision tree type of diagnostic evaluation with the plan mimicking a flow chart. The execution of this plan can involve ordering diagnostic tests, and making clinical observations with the patient physically and emotionally. Using that input, care strategies can be followed according to treatment and pathway guidelines to improve the patients' health. Disease management has always been important, but electronic health records and data analysis have provided much more information to enable better patient

care decisions. Regional Extension Centers, evidence-based medicine, and care coordination are examples of disease management.

Disease management is an approach to health care that teaches patients how to manage a chronic disease (Tory 2020). Self-management is a means to disease management. Self-management involves providing patients with education and tools to empower them to manage and improve their own condition and improve their quality of life. This is used primarily in cases of chronic conditions such as diabetes, hypertension, heart disease, renal disease, cancer, arthritis, asthma, and lung diseases.

A study that has been successful in utilizing this approach and successfully providing data analysis to reflect outcomes is the Stanford Chronic Disease Self-Management Program (CDSMP). The most impressive outcomes of the study are that there was "significant, measurable improvements in the health and quality of life of the adults with chronic conditions" with savings "enough through reductions in healthcare expenditures to pay for itself within the first year" (NCOA 2015). This program illustrates how patients that are trained to handle their chronic diseases through better self-management of their conditions, can improve their health status, and reduce their need for more costly medical care.

Disease registries are another popular way to manage patients with chronic diseases. Disease registries group patients together so that clinicians can evaluate and improve outcomes for a population that is defined by a certain disease. There are tumor registries, cancer registries, Alzheimer's registries, etcetera all with the focus on aggregating large data sets to analyze trends or patterns in treatments and outcomes. Registries can also be used for quality assessments, clinical improvements, benchmarking, cost effectiveness, clinical research, population surveillance, and treatment surveillance (Moore 2021). There is a very comprehensive list of registries available at NIH where patients that have particular diseases and might be eligible for new studies or drug trials (NIH 2022).

Regional Extension Centers

Regional Extension Centers (RECs) serve as a support and resource center to assist providers in EHR implementation and Health IT needs. In an advisory capacity, the RECs "bridge the technology gap" by helping providers navigate the EHR adoption process from vendor selection and workflow analysis to implementation and meaningful use (HHS 2020). The US Department of Health and Human Services established three stages to measure use of EHRs in a "meaningful manner" with Stage 1 of the process to measure use of health information that focuses on acquiring a baseline of information on patients. This process supports measurement of the expected outcome of an improvement in the quality of care provided, as well as promoting efficiencies in care delivery. The funding for the development of RECs came from the Health Information Technology for Economic and Clinical Health Act (HITECH). REC activities include the enabling of HIE for patient-centered health IT, and patient engagement while maintaining privacy and security. An example of a program developed that relies heavily on physicians utilizing technology to monitor patients is the Million Hearts Program. This is a national initiative whose goal is to prevent one million heart attacks and strokes by 2027 (HHS 2022). A key component of this initiative is that patient information from the physicians' offices is monitored and appropriate interventions are taken when defined triggers are identified. These interventions can be improving patients' access to care, smoking cessation programs, cholesterol management, promoting a healthy lifestyle and educational resources. Every 36 seconds in the United States, an adult dies from cardiovascular disease (CVD). This program which started in 2012 is aimed at averting one million preventable CVD events in the next five years (January 2022–December 2026). Their

program can include decreasing tobacco use, monitoring blood pressure, decrease inactivity and monitoring other health related items to achieve their goal.

Clinical Decision Support

The advances and utilization of technology in healthcare have provided many ways to improve healthcare. One functionality that relies on the EHR is clinical decision support (CDS), which "provides persons involved in care processes with general and person-specific information, intelligently filtered and organized, at appropriate times, to enhance health and health care." (ECQM 2022) A traditional CDS is comprised of software designed to be a direct aid to clinical-decision making, in which the characteristics of an individual patient are matched to a computerized clinical knowledge base and patient-specific assessments, or recommendations are then presented to the clinician for a decision. (NPJ 2020) CDS systems are basically implementing evidence-based clinical guidelines at the point of care. Advantages to CDS are the reduction in the number of errors and diverse events (drug administration errors, wrong site surgery and other surgical injuries, patient falls), improved efficiency, decreased costs in providing care, and better provider and patient satisfaction (HHS 2016). With a good clinical decision support system built into the EHR, clinicians can benefit from behind-the-scenes suggestions that are autogenerated. Use of knowledge bases of medical facts, disease guidelines and other reference data can be coupled with the specific patient data. This data can then be used as input to an inference engine or algorithms designed to give specific advice or recommendation around the treatment of the patient. There are certain patterns of care that are standard for many chronic diseases; suggestions and checks can be done in the background to limit omissions in testing or by adding alerts and reminders for clinicians. It is important to note that adding too many alerts and reminders can cause "alert fatigue" which can have the opposite effect, providers ignore the alerts. (Alert Fatigue 2021).

Evidence-Based Medicine

Evidence-based medicine (EBM) is "the conscientious, explicit, judicious and reasonable use of modern, best evidence in making decisions about the care of individual patients" (NIH 2021). EBM's goal is use clinical knowledge, patient's desires and the best research and establish optimal outcomes for patients. The strongest evidence to support and aid in making medical care decisions is established by assimilating clinicians' expertise, patient values, and desires with clinically sound, relevant research. For clinicians EBM implies being aware of treatment options, research and making decisions based on clinical need and the effectiveness of treatment options. Clinicians making EBM decisions should assess evidence quality and reproducibility, and discuss evidence and recommendations with the patient.

Seamless Care

The definition of seamless care or care coordination is organizing patient care activities and sharing information among all participants concerned with the patient well-being to achieve safer and more effective care (AHRQ 2018). Patients who receive continuity of care have better healthcare outcomes, and higher satisfaction rates. Patient connectivity with healthcare providers can be through telemedicine follow ups, visiting nurses, patient portals and communication among the healthcare provider team. The focus of seamless care is always on the patient. One common thought emerging in most discussions of health informatics systems to support seamless care is that much of the information needed is the same data required in direct patient care. Many times patient care data is in the EHR or siloed so that the entire healthcare team does not have access to it. Lack of information, even minor details can be omitting a clue that will allow a better diagnosis.

The redesign of systems is necessary to make systems more patient-centric, where all health data about a patient are aggregated to form a single patient EHR. Systems for maintaining patient information, even when EHR systems are used, often fail to support continuity of care due to multiple providers, lack of interoperability between systems, lack of standards for systems, and lack of standardized definitions and terminologies. These are issues that are directly related to the lack of optimal utilization of health information ICS including the lack of multiple providers, interoperability, and standards. The CMS Innovation Center's Comprehensive End Stage Renal Disease (ESRD) Care Model exemplifies seamless care in the current healthcare environment. The participants are referred to as ESRD Seamless Care Organizations (ESCOs). The high cost of treating ESRD patients due primarily to the underlying complications and comorbidities is the primary focus of the program. There are 600,000 patients in the United States receiving regular dialysis treatments with 6.3 percent of Medicare spending of over $30.0 billion for ESRD patients (CMS 2016). This model provides an example of how care coordination can improve care and help control costs of care by:

page 469

- Implementing continuous data-driven learning
- Improving quality of life and functional status
- Coordinating clinical practices for efficiency and improved quality care
- Standardizing performance metrics and benchmark measures
- Minimizing unnecessary ED visits
- Enhancing and coordinating patient-centered care
- Improving beneficiary access to services
- Reducing dialysis-related complications
- Promoting freedom to seek services and providers of choice. (CMS 2016)

More than one in seven adults in the United States (15%) or 37 million people, are estimated to have chronic kidney disease; with these staggering numbers any improvements in treatments or models will have a big benefit to both patients and cost containment.

Implementing Emerging Technology in Health and Healthcare

After discussing topics surrounding the emerging technologies themselves, the next step is to explore cases of implementing the technologies that are creating advancements in health platforms and healthcare. Technology has been increasing rapidly and its application to healthcare has naturally followed. The following sections will explore two cases of implementation of emerging technology, specifically in the use of artificial intelligence and cybersecurity. New medicines and vaccines are fundamental aspects of our current healthcare systems and for improvement of public health. Use of AI to determine molecular structures can be a critical aid in designing new drugs and treatments, and there have been recent breakthroughs. Healthcare institutions spend a great deal of time and effort on collecting patient data from EHRs and from many other sources. Protecting that data from malware and other threats has become important to the patients' well-being and an institutions survival. The patient history tells a detailed story about the patient and is important for future care and must be guarded. Some implementation examples of these technologies follow.

page 470

Case of Implementation: AI and AlphaFold

For more than a century it has been possible to accurately determine the atomic structure of large and small molecules. Proteins, DNA, substrates, and viruses, because of their size, were more of a challenge, but many structures are known through various techniques. The most accurate technique for this use is x-ray crystallography, where atomic distance is known to very exact measurements in Angstroms (10^{-10} m). Determining the atomic structure of DNA, proteins, and viruses using these techniques is a very difficult and lengthy process. Proteins are complex molecules and do most of the work in cells. They go from a string of amino acids to a machine that drives life. They are important to the structure, function, and regulation of many processes in the body. Knowing a molecular structure like a protein provides a greater level of understanding of how a process works. This can allow formulation of hypotheses about how to affect it, control it, or modify it. It could give critical insight into the nature of a disease and what might be done to stop or slow its progression. Proteins are involved in the immune system, can allow organs to communicate between each other, and repair damaged tissue, to name just a few of their functions. Protein structures are known to affect diseases like Alzheimer's, Type 2 diabetes, ALS, and Parkinson's disease. Knowing a protein's structure could allow the design of site-directed mutations with the intent of changing function or identifying active sites. This knowledge could be used to help determine the function and mechanisms of action, which can lead to design of molecules to adhere to and block the active site, depending on the desired effect.

AlphaFold2 was created by a Google subsidiary called DeepMind, which is a group of AI and machine learning (ML) researchers. They previously had been focused on solving the challenge of general intelligence, that is, creating an AI that could solve any problem. In the journey to the general intelligence holy grail, Deep Mind started working on solving smaller problems. First, they developed an algorithm to play Atari games using reinforcement learning, where the AI had to learn Atari games by trying to maximize the score. After algorithm development, training and refinement, the AI was unbeatable at almost all games. Deep Mind then went on to conquer the ancient Chinese game of Go where after several iterations was able to beat the world champion Lee Sedol on March 19, 2016. In a more scientific endeavor Deep Mind choose to predict protein structures. Since its start in 1971, the Protein Data Bank archive (PDB) has served as the single repository of information about the 3D structures of proteins, nucleic acids, and complex assemblies in the world (PDB 2022). These known protein structures were used to train the Deep Mind AI application. With Alpha Fold an amino acid sequence is used as input, and the result is a 3D dimensional model based on complex interrelationships. The algorithms that were used for Alpha Fold were harnessing convolutional neural networks (CNN) that is a deep learning approach to structure prediction. A CNN is a form of neural networking that are very good at picking up on patterns in the input. The architecture of a convolutional neural network is a multi-layered feed-forward neural network, made by stacking many hidden neuron layers on top of each other in sequence. Adjusting these hidden layers with an iterative weighting scheme which creates a predictive model. It is this sequential design and iterations that allows convolutional neural networks to learn hierarchical features. Deep Mind researchers coupled these algorithms with other algorithms that could predict structure alignment of segments of the protein. Demis Hassabis has been the driving force behind DeepMind and has said "It's been so inspiring to see the myriad ways the research community has taken to AlphaFold, using it for everything from understanding diseases, to protecting honey bees, to deciphering biological puzzles, to looking deeper into the origins of life itself." (DeepMind 2022).

This work can have a profound effect on healthcare in the future by providing new medicines that can block active disease sites or alter proteins which could potentially

change the conformation of a molecule that might be responsible disease.

Ethical Concerns

Ethical issues must be considered with the data generated through Health Informatics applications. These considerations include the areas of autonomy, beneficence, nonmaleficence, and justice. Because there is an increased use of ML and AI in healthcare, there are growing ethical concerns about how algorithms are trained. Should an AI or ML project be required to disclose the algorithms used and data that were used to refine and train the algorithm? Are there intrinsic biases in the data that will lead to erroneous output? For diagnostic algorithms where AI might be examining cancer cells or looking for patterns in data, it is very important to look for any biases that might be built into a model. There should also be disclosure of how many samples were used to train a model. This is all pertinent information considering important clinical decisions can be largely influenced by computer-generated output. Although AI and ML tools are growing in popularity, they still require human review, in keeping with the "do no harm" tenet. There is a duty for each clinician to have confidence and assurance in the safety of using new tools and the algorithms on which they are based.

Case of Implementation: Cybersecurity

In 2021, 45 million individuals were affected by healthcare cyberattacks, up from 34 million in 2020. That number has tripled in just three years, growing from 14 million in 2018, according to the report, which analyzes breach data reported to the HHS by healthcare organizations (Landi 2022). Many of these threats are very disruptive to patient care, whether it is ransomware, disruption in the medical supply chain, or other exploits, cybercriminals want to cash in by selling protected health information (PHI) and account information on the dark web.

Cybersecurity can be defined as "precautions taken to guard against crime that involves the internet, especially unauthorized access to computer systems and data connected to the internet" (Dictionary 2022). Good cybersecurity is a combination of people, security processes, and technologies. Regular antivirus scans, both on data at rest and data in motion, are important. Having firewalls with rules that are regularly updated will also help minimize attack points. Another good security practice is to hire ethical hackers, sometimes referred to as white hats, to try to penetrate an EHR or analytic system. Having security experts look for common or uncommon weaknesses before cybercriminals find them can help an institution avoid an attack.

With health information access, specifically to the EHR system and anywhere else PHI is stored, cybersecurity is critical. An executive order was issued in 2013 by President Barack Obama that called for the development of a cybersecurity framework to reduce and manage cybersecurity risks (HealthIT 2016). President Joe Biden issued a similar executive order for cybersecurity for the detection, assessment, and remediation of cyberincidents (Biden 2021). A number of groups are actively involved in cybersecurity initiatives. NIST published the Framework for Improving Critical Infrastructure Cybersecurity in support of providing a more secure infrastructure using a set of activities to achieve specific cybersecurity outcomes (NIST 2021). The Office of the National Coordinator for Health Information Technology (ONC) has been involved with funding initiatives and continues to develop educational resources for healthcare cybersecurity and risk management among other activities (ONC 2016). As one such activity, the ONC has laid out a Federal Healthcare IT Strategic Plan for 2020-2025 (ONC 2020). The ONC recommends for organizations to have an appropriate protection plan in place. The following suggestions are included in this plan:

- Establish a security culture.

- Protect mobile devices.
- Maintain good computer habits.
- Use a firewall.
- Install and maintain antivirus software.
- Plan for the unexpected.
- Control access to PHI.
- Use strong passwords and change them regularly (two-factor ID preferred).
- Limit network access.
- Control physical access.
- Train staff on good security practices.

These are standard operating procedures that most individuals and institutions are familiar with, but they are the basic requirements for protection from cybersecurity attacks. As a first step, top management must create a culture that supports a secure workplace. They must establish policies and procedures for accessing, analyzing, modifying, storing, transmitting, and sharing data. Today's workplace is collaborative and can be geographically separated, particularly since the COVID-19 pandemic, as laptops have become extensions of an institution's computing architecture. Workers may access data remotely from home or while traveling. They might use their work computers or a personal device such as a mobile application on a cell phone. Similarly, data may be entered into business systems from unsecured personal devices or software on the users' personal computers. These activities should be restricted and only approved applications should be allowed to access healthcare systems containing PHI data. Failing to establish best practices for protecting PHI is an invitation to hackers. As laptops now can access the deepest reaches of an institution's data, they must be more than casually protected. Common methods are using two-factor identification and whole disk encryption so that if a laptop is stolen, data are not visible. The following are best practices for protecting PHI.

- Ensure mobile devices are equipped with strong authentication and access controls.
  - Ensure laptops have password protection or two-factor log on.
  - Enable password protection on handheld devices (if available). Take extra physical control precautions over the device if password protection is not provided.
- Protect and encrypt wireless transmissions from intrusion.
- Do not transmit unencrypted PHI across public networks (internet, WiFi).
- Where it is absolutely necessary to store PHI on a mobile device or remove a device from a secure area, encrypt the data.
- Do not use mobile devices that cannot support encryption.
- Develop and enforce policies specifying the circumstances under which devices may be removed from the facility.
- Take extra care to prevent unauthorized viewing of the PHI displayed on a mobile device.
- Run intrusion detection software to monitor unusual behavior trying to access your system.
- Use an antivirus product that provides continuously updated protection against viruses, malware, and other code that can attack computers through web downloads, CDs, email, and flash drives.

- Keep antivirus software up to date. Most antivirus software automatically generates reminders about these updates, and many are configurable to allow for automated updating. (HHS 2010, 27, 41)

Check Your Understanding 16.1

Answer the following questions.

1. Which of the following is a historical copy saved at a specific point in time? page 474
   a. Backup copy
   b. The cloud
   c. Archival system
   d. Master data
2. Which one of the following is an example of POCT?
   a. A patient taken to the x-ray department with the report then available in the EHR system.
   b. Previous x-ray reports from other facilities are sent to the hospital to be scanned and are available in the patient's record
   c. Portable x-ray equipment is taken to the patient's room where the x-ray is taken, with a report later available in the patient's EHR
   d. A patient is referred to an outpatient radiology center to have an x-ray done at their convenience.
3. Accountable care organizations:
   a. Support healthcare reform and implementation of the EHR
   b. Offer care that is consistent and of uniform quality
   c. Are a group providing care at the instance of illness and injury
   d. Are a group of providers with the responsibility for improving health status, efficiencies of care, and experiences for a defined population
4. The process for bringing together and summarizing large amounts of data from multiple sources is:
   a. Data consolidation
   b. Purging
   c. Data manipulation
   d. Developing a data dictionary

Provide a short answer for the following.

5. Tailoring treatment to the medical characteristics of the individual patient is:
6. The challenge for storing large amounts of clinical data from many institutions is due to:

Health Informatics Workforce

The field of HI continues to progress with the development of new technologies, as new uses for healthcare data are identified and regulations change. "The demand for health informatics workers is projected to grow at twice the rate of employment overall, but there is strong evidence that the nation already faces a shortage of qualified workers in this field" (Burning Glass 2014). According to the Bureau of Labor Statistics about 34,300 openings for medical records and health information specialists are projected each year, on average, over the decade. Many of those openings are expected to result from the need to replace workers who transfer to different occupations or exit the labor force, such

as to retire. (BLS 2022). As the volume of health data grows, the need to report, analyze, and visualize data grows as well. The need for HI professionals to handle this growth must align with the growth; otherwise, important data and findings could go unexplored. Educational requirements and workforce aspects are addressed next.

Education

Formal education in HI has been available for a number of years either as degree programs or courses within other programs. These programs, especially at the master's degree level, have proliferated over the past few years to coincide with the increased use of EHRs and recognition of the value of the data now available. The curriculum varies among the programs and continues to change with the continued evolution of the field. Although the definitions of biomedical, clinical, nursing, and healthcare informatics, among others, continue to be debated, commonly accepted definitions of the terms are provided here. Biomedical is related to both biology and medicine. A clinical nurse specialist is a licensed registered nurse with graduate preparation at the master's or doctoral level and has expert clinical knowledge and skills in a specialized area of nursing practice (National Association of Clinical Nurse Specialists 2016). Healthcare or medical informatics is defined as an interdisciplinary study of the development and application of IT in the planning, delivery, and management of healthcare services (HIMSS 2016). IT personnel cannot deal effectively with medical semantics, yet to clean and homogenize data for research, reporting, and analysis this is a vital task. Therefore, HI really requires a tight bonding of both clinical and IT personnel to be effective. What the field should become is of more significance with program curriculum development activities conducted to meet the current demands to develop an informatics workforce, as well as the needs of the future.

Educational Programs

The field of HI is broad. There are a number of different subject matter areas, and different job positions have been created to do the work of informatics. Biomedical and HI has been purported to be "the most comprehensive term for all fields concerned with the optimal use of information, often aided by the use of technology to improve individual health, healthcare, public health, and biomedical research" (Hersh 2009, 99). Studies have been conducted to compare the curricula of a variety of programs across the US and to compare courses within these curricula.

A recent review was made of the websites of current HI degree programs throughout the United States. The focus and the curricula of these programs continues to be varied. The majority of the degree programs reviewed were at the master's level with program names such as Health Informatics, Health Information and Informatics, Medical Informatics, Bioinformatics, and Nursing Informatics. All of these programs were found to be fairly consistent, with the more common content areas noted above, but there were differences that provided a unique identity and focus to each individual program. The focus areas that distinguished one program from another are combining aspects of health information with HI, applications and use of technology, research focus, incorporation of biomedical informatics courses with medical school courses, policy development, incorporation of engineering with the study of biomedical informatics, and advanced practice for nursing leadership.

The Commission on Accreditation for Health Informatics and Information Management Education (CAHIIM) is an accrediting organization whose mission is to serve the public interest by establishing and enforcing quality accreditation standards for HI and information management educational programs (CAHIIM 2022). They are the accrediting body for health information management associate's, bachelor's, and master's

degree programs, which provide the pathway for eligibility for the Registered Health Information Technician (RHIT) and Registered Health Information Administrator (RHIA) certifications. One of the six domains in the curriculum requirements for these programs is Informatics, Analytics, and Data Use (Domain III).

CAHIIM also accredits HI programs at the master's degree level. As published by CAHIIM, HI graduate programs focus on information systems, informatics principles, and IT as applied to the continuum of healthcare delivery. Individual programs offer options for areas of practice or research (CAHIIM 2022).

There are a number of HI degree programs in the US that have been developed independent of the CAHIIM accreditation guidelines. These programs tend to be developed to meet a specific need in the workforce or to meet the interests and expertise of the faculty. The majority of these degree programs are at the master's and doctoral levels in biomedical or clinical informatics. In addition to meeting a particular workforce need, other critical goals of these educational programs are to promote research in the HI domain to validate the profession and define the body of knowledge, develop best practices, and find new applications in healthcare for the potential outcomes related to the intersection of information and technology.

Certificate programs are short-term, nondegree programs and are typically focused on a narrow subject area. They can provide new opportunities for employment and are usually developed to meet a specific need in the workforce. There are certificate programs designed to address assorted aspects of HI. Some of these are offered by colleges and universities in addition to their degree programs and some are offered by professional associations. Others were developed with funding from the ONC for workforce development. Certifications may be earned by taking an examination by an accrediting or approval agency.

AHIMA offers of a number of credentialing examinations, with three that are consistent with the HI domains. The Certified Healthcare Technology Specialist (CHTS) credentialing exam is associated with various roles related to the implementation and management of EHR systems. The Certified Health Data Analyst (CHDA) credential supports better use of information. Another certification is the Certified Professional in Health Informatics (CPHI) to recognize the knowledge to be able to provide support for and the ability to utilize HI (AHIMA 2016).

Table 16.1 . Informatics and data analysis potential job titles with associated skills required

Position	Specific skills required for position
Chief Clinical Informatics Officer Medical Informatics Officer Related Jobs: Professor Consultant	• Managerial • Technical and professional skills of systems/procedure analysis • Programming and hardware capabilities • Strong skills in performance management, quality, and data analytics • Knowledge of EMR Applications Systems • High level of problem-solving skills • Ability to work and communicate effectively with all levels of management, both inside and outside the Information Management division • Experience and judgment to plan and accomplish goals • Ability to work in a dynamic environment

Director of Clinical Informatics Director of Medical Informatics Related jobs: Program Director Consultant	• Organizational skills at the director level • Leadership skills with good clinical, physician, and executive communication • Clinical experience • Formal or informal leader within the facility • Listens carefully to respond appropriately, answer questions, or obtain needed information • Demonstrates flexibility in adjusting to fluctuating needs • Manages the time and availability of personnel and other resources effectively • Proven experience in troubleshooting and problem solving in a demanding management environment • Thorough working knowledge of the facility processes and personnel
Research and Development Specialist Clinical Research Informatics, Lead Analyst Related Jobs: Chief Learning Officer Professor Program Director	• Experience and knowledge in health IT • SQL, Python or other programming skills, Terminologies skills • Cloud Knowledge • Experience in Data Visualizations
Project Manager Related Jobs: Director of Clinical Informatics Business Analyst	• Strong presentation skills • Experience in MMIS or other healthcare claims system(s) IT projects preferred • Skills/experience with Medicaid or Healthcare Delivery Systems preferred • Experience working with SharePoint and other typical project management tools, including Visio, PowerPoint, Word, Access, and Excel

Mapping Specialist Point-of-Care Mapping Specialist Related Jobs: Coding Trainer Coding Manager Consultant Director of Clinical Informatics	• Understand and apply rule-based mapping guidance • Follow established data mapping principles and published best practices • Possess knowledge and/or experience in the source and target systems involved in the map • Demonstrate diversified coding and data management experience with clinical terminologies • Have knowledge of current global healthcare informatics standards • Mastery of medical terminology and medical science principles • Broad understanding and experience with healthcare vocabularies including the World Health Organization's Family of Classifications • Health informatics training or equivalent experience required • Clinical information system experience desirable • Ability to acquire knowledge and skill in information retrieval for map development and use of mapping tools and cloud processing • Ability to meet required delivery dates
Data Integrity Analyst Data Integrity Officer Related Jobs: Consultant Data Quality Manager	• Must have the ability to analyze healthcare data from conceptualization through presentation of the data • Proficiency with analytical tools (SQL, Excel, Python, R) • Knowledge of data analysis methodology • Skill with presentation software • Strong communication skills • Cloud skills for scaling and processing
Informaticist Researcher Applied Health Informatics Researcher Related Jobs: Clinical Informatics Coordinator	• Clear understanding of research design and how to develop accurate and correct questions for surveys and research projects • Ability to work in a team environment and with those familiar with research and those not as familiar is essential • Excellent computer skills and knowledge of databases and software programs available for research is very important • Understanding of medical procedures is critical for designing effective research projects • Knowledge of hospital, state, and federal regulations regarding medical procedures and policies as well as data collection and confidentiality is a must for an informatics researcher

Clinical Informatics Coordinator Clinical Information Systems Coordinator Related Jobs: Director of Clinical Informatics Informaticist Researcher	• Working knowledge of computers and computer systems as well as hands-on experience working in a clinical setting • Experience working with various groups • Patience to teach computer skills to a wide variety of individuals • Experience with advanced clinical software systems, databases, and other computer software programs developed for hospital or healthcare facilities • Additional training in computer skills and IT is strongly encouraged
Data Analyst Data Informatics Analyst Related Jobs: Informaticist Researcher Data Integrity Analyst	• ICD mapping • Knowledge of CPT, Terminologies • SQL, Python, SAS, R skills • Knowledge of cloud and scalable processing • Visualization Skills
Content Analyst Related Jobs: Informaticist Researcher Data Architect	• Negotiation • Persuasion • Evaluating requirements to determine impact to resources, schedules, tools, and process

There are other organizations that offer opportunities for certification. The American Medical Informatics Association (AMIA) approved the Advanced Health Informatics Certification (AHIC) in 2019. Once an individual meets the AHIC eligibility criteria, recertification requirements, and passed the exam, they are known as an AMIA Certified Health Informatics Professional (ACHIP). The American Society of Health Informatics Managers offers an exam to earn the Certified Health Informatics Systems Professional (CHISP) credential. The certification exam offered by Health Information and Management Systems Society (HIMSS) is to earn the credential, Certified Professional in Health Information and Management Systems (CPHIMS).

Demand

Reports, research, visualizations, and analysis in medicine indicate a tremendous need for HI workers. It is sometimes difficult to determine what the positions are because they can range from data analysts to chief information officers (CIOs) or informatics researchers. The adoption of the EHR and the increased reliance on data and data analytics are given as primary reasons for the high demand for workers at all levels, with differences in the required knowledge, skills, and experience. Specific statements regarding the openings are as follows. More than 50,000 HI openings were projected through 2020 to 2022. Demand is expected to increase at twice the rate of all healthcare jobs overall, along with the fact that there is already a shortage of HI workers (Burning Glass 2014). Informatics jobs tend to remain open longer than other types of jobs, indicating the difficulty in filling these positions. HI positions require a more varied skill set for the informatics employees to have the ability to remain current with the continued evolution of regulations and the varied needs of big data processing such as the increased utilization, complexity, and distributed nature of the information. Some positions can be considered hybrid, requiring

skills in more than one discipline, such as health information management (HIM) and computer technology or nursing and data analysis. A review of job postings reveals a great deal of diversity in the positions that might be available based on job title. For example, a brief list of some of these might include the following:

- Business Intelligence Analyst
- CHIO, Physician
- Clinical Information Liaison
- Clinical Information Specialist
- Coding Validator
- Data Analyst
- Data Scientist
- Data Visualization Specialist
- Health Informatics Director
- Health Informatics Faculty
- Health Informatics Scientist
- Health Information Solutions Coordinator
- Health Information Technician
- Health Informatics Project Manager
- Project Analyst
- Public Health Scientist
- Research Analyst
- Senior Consultant, Population Health
- Senior Research Analyst
- Systems Support
- Systems Support Trainer

Positions and Skills Needed

The diversity of the kinds of health informatics positions results in the possibility that different knowledge and skills are needed for individual jobs. Table 16.1 presents a list of identified informatics-related positions from the AHIMA Career Map (AHIMA 2022). The Career Map illustrates positions by level and is an interactive site where career transition and promotion opportunities can be explored. The AHIMA Career Map site provides additional information about each of the positions listed there to include a description of the job, and the responsibilities and skills required. See table 16.1 for a listing of the positions illustrated at Career Map in the section on Informatics and Data Analysis and the specific skills required for each of these positions. In addition to these job specific skills, a number of skills emerged as important for each of the jobs. These general skills for the informatics/data analysis jobs are: working knowledge of Microsoft Suite, detail orientation, independent, excellent interpersonal oral and written communication skills, basic programming ability (such as python and R), ability to multitask, problem solver and troubleshooter, critical thinking skills, time management, ability to function in a leadership role, organizational skills, ability to interact with professionals at all levels with diplomacy and tact.

Check Your Understanding 16.2

Answer the following questions.

1. To enter the health informatics field, an individual need to be mindful of the following statements, except:
   a. Visualization skills are important to show data in concise way and allowing decision makers to explore the data easily and conveniently.
   b. All health informatics jobs require a degree in informatics.
   c. Basic programming ability (such as Python and R) would help in many situations as health informaticists deal with big amount of data.
   d. Understanding the cost and the opportunity of cloud technology will help health informaticists determine whether to host the technology on the cloud or on-premise.
   e. Ability to interact with professionals at all levels with diplomacy and tact is a must since health informaticists bridge the communication between technology and practitioners.

2. Why does health informatics jobs remain open longer than many other posted jobs?
   a. The varied skill set that might be required for the job.
   b. The limited number of position available.
   c. Lack of interest in joining health informatics.
   d. Low incentives of joining health informatics due to lower salary.
   e. Healthcare professionals prefers more tasks that directly works with patients.
3. Which is not a generally required skill across most health informatics position?
   a. A working knowledge of ICD-10 coding (or some type of coding system)
   b. Knowledge of the EHR
   c. In-depth hands-on clinical expertise
   d. Visualization skills
   e. Negotiation and communication skills
4. Why is the AHIMA Career Map important?
   a. It is a fancy way to map health need.
   b. It can be used to obtain information on some of the emerging positions with required job skills and advancement opportunities.
   c. I can show it to my colleagues.
   d. It offers competitive advantage in negotiating for salary and career progression.
   e. It allows individuals to think about going back to the education process.
5. Cloud skills are becoming increasingly important in medical data analysis. What do you need to understand about cloud technology?
   a. Browser knowledge, computing technology knowledge, and user interface knowledge
   b. Network knowledge, return of investment from the Internet, and advanced machine learning algorithm
   c. Storage knowledge, network knowledge, and programming knowledge
   d. Network knowledge, computing technology knowledge, return of investment from cloud technology

---

Use of technology and data drive real progress in healthcare. Data live in so many places, the EHR, Medical devices (e.g., Fitbit, Smart Blood Pressure cuff), Lab instruments, specialized systems (e.g.GI, Cardiac, Blood Bank etc.), patient portals and more. Combining this data so that it is patient centric can help clinicians "see" a full view of the patient. This chapter has described many topics related to disease management, quality improvements, evidence-based medicine, accountable care organization and many

other techniques and approaches all focusing on better healthcare outcomes and controlling costs. To understand and apply current advances and discoveries, it is necessary to have a firm grasp of healthcare informatics. There are many healthcare websites both commercial and government, journals, books and other resources to maintain and enhance your education. Knowledge from formal and informal education is the foundation that can drive and influence new solutions or help make improvements in early disease detection, treatments and guidelines leading to better patient outcomes. It is important for a healthcare informatics specialist to be able to discern what technologies could be applied to specific challenges and be able to problem solve choosing the right tools and correctly estimating the required resources to be successful.

Chapter 16 Review Exercises

Answer the following questions.

1. Which of the following is all the genetic information present in an organism?
    a. Genomics
    b. Genome
    c. Genetics
    d. Functional genomics
2. What began with the identification of the human genes and led to the development of tools for data analysis?
    a. Genome
    b. Genetics
    c. Human Genome Project
    d. Genomic informatics
3. Which is a historical copy of data saved in another country?
    a. Backup
    b. Archival database
    c. Offshore data storage
    d. The cloud
4. Which is an accrediting organization that would assess compliance with accreditation standards for a master's of health informatics program?
    a. AHIMA
    b. CAHIIM
    c. HIMSS
    d. AMIA

Provide a short answer for the following.

5. Which nationally recognized organization developed a framework for cybersecurity infrastructure?
6. Which of the following is the state of being protected against the criminal and unauthorized use of data?

Discussion

7. What is the difference between traditional health informatics programs and health informatics certificate programs?
8. What is the aim of AlphaFold?
9. List at least three emerging technologies in health informatics.

REFERENCES

Adshead, Anthony Snapshots 101: Array vs backup software. https://www.computerweekly.com/feature/Snapshots-101-Array-vs-backup-software.

Alert Fatigue. 2021. ***EHR Intelligence***. "Best Practices to Mitigate EHR Clinical Decision Alert Fatigue". https://ehrintelligence.com/news/best-practices-to-mitigate-ehr-clinical-decision-alert-fatigue.

American Health Information Management Association (AHIMA). 2016. Types of Certifications, Specialty. https://www.ahima.org/certification-careers/certification-exams/.

American Health Information Management Association (AHIMA). 2017. ***Pocket Glossary for Health Information Management and Technology***. 5th ed. AHIMA Press.

American Health Information Management Association (AHIMA). 2022. Health Information Careers, Career Map. https://my.ahima.org/careermap.

Agency for Healthcare Research and Quality (AHRQ). 2018. Care Coordination https://www.ahrq.gov/ncepcr/care/coordination.html#:~:text=Care%20coordination%20involves%20deliberately%20organizing,safer%20and%20more%20effective%20care .

Agency for Healthcare Research and Quality (AHRQ). 2021. Defining the PCMH. AHRQ Patient Centered Medical Home Resource Center. https://www.ahrq.gov/ncepcr/tools/pcmh/defining/index.html.

AHRQ. 2014. AHRQ Health Care Innovation Exchange. Trends in Health Information Exchanges. https://innovations.ahrq.gov/perspectives/trends-health-information-exchanges.

Alzu'bi, A., L. Zhou, and V. Watzlaf. 2014. Personal genomic information management and personalized medicine: Challenges, current solutions, and roles of HIM professionals. ***Perspectives in Health Information Management*** 11(Spring):1–22.

Amazon AWS. n.d. "What is a Data Lake" https://aws.amazon.com/big-data/datalakes-and-analytics/what-is-a-data-lake/. Accessed August 13, 2022.

An, G., and Y. Vodovotz. 2016. What is "Precision Medicine"—and can it work? Elsevier Connect. https://www.elsevier.com/connect/what-is-precision-medicine-and-can-it-work.

Biden. 2021. Executive Order on Improving the Nation's Cybersecurity. https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity/ .

Bonderud, D. 2021 ***Health Tech Magazine.net*** April 2021. "How Predictive Modeling in Healthcare Boosts Patient Care" https://healthtechmagazine.net/article/2021/04/how-predictive-modeling-healthcare-boosts-patient-care-perfcon .

Bureau of Labor and Statistics (BLS). 2022. "Medical Records and Health Information Specialists" https://www.bls.gov/ooh/healthcare/medical-records-and-health-information-technicians.htm.

Burning Glass Careers in Focus. 2014. ***Health Informatics 2014: Missed Opportunities? The Labor Market in Health Informatics, 2014***. Burning Glass Technologies 2014 Report. http://burning-glass.com/research/health-informatics-2014/.

Centers for Disease Control and Prevention (CDC). 2016. Genomics and Precision Medicine. https://www.cdc.gov/genomics/gtesting/index.htm.

Centers for Disease Control and Prevention (CDC). n.d.a., "Epidemic Intelligence Service". https://www.cdc.gov/eis/index.html. Accessed August 11, 2022.

Centers for Disease Control and Prevention (CDC). n.d.b. Health care provider/clinician apps. http://www.cdc.gov/mobile/healthcareproviderapps.html. Accessed August 11, 2022.

Centers for Disease Control and Prevention (CDC). n.d.c. "Public Health Grand Rounds". https://www.cdc.gov/grand-rounds/index.html. Accessed August 13, 2022.

Centers for Disease Control and Prevention (CDC). n.d.d. "COVID Data Tracker". https://covid.cdc.gov/covid-data-tracker/#datatracker-home. Accessed August 13, 2022.

Centers for Disease Control and Prevention (CDC). n.d.e. "Patient-Centered Medical Home (PCMH) Model". https://www.cdc.gov/dhdsp/policy_resources/pcmh.htm. Accessed August 14, 2022.

Centers for Disease Control and Prevention (CDC). n.d.f. "National Center for Chronic Disease Prevention and Health Promotion (NCCDPHP)". https://www.cdc.gov/chronicdisease/resources/publications/aag/population-health.htm. Reviewed July 13, 2021.

Center for Medicare and Medicaid Services (CMS). 2016. Comprehensive ESRD Care Model: RFA Fact Sheet. CMS. https://innovation.cms.gov/Files/fact-sheet/cec-py2.pdf.

Center for Medicare and Medicaid Services (CMS). n.d.a. Accountable Care Organizations (ACOs). https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/ACO. Accessed: August 14, 2022.

Center for Medicare and Medicaid Services (CMS). n.d.b. CMS Redesigns Accountable Care Organization Model to Provide Better Care for People with Traditional Medicare. https://www.cms.gov/newsroom/press-releases/cms-redesigns-accountable-care-organization-model-provide-better-care-people-traditional-medicare and https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/ACO. Reviewed February 24, 2022.

Chou, Y. 2022. Gamification and Behavioral Design. "Top Ten Gamified Healthcare Games that will extend your Life" Accessed August 13, 2022. https://yukaichou.com/gamification-examples/top-ten-gamification-healthcare-games/.

Commission on the Accreditation of Health Informatics and Information Management Education (CAHIIM). 2022. Health Informatics. https://cahiim.org/hi/hi.html.

CopperMobile. 2022. Aakash Sareen. March 7, 2022. "What you should know about Gamification in Healthcare in 2022". https://www.coppermobile.com/blog/gamification-in-healthcare-2022/.

Critical Insight. 2021. Healthcare Breach Report July – Dec 2021. Security Research and Data Analysis – Critical Insight. www.criticalinsight.com.

Dictionary. 2022. Cybersecurity Definition. Dictionary.Com. Accessed August 15, 2022. https://www.dictionary.com/browse/cybersecurity.

Daily Advisory. 2018. ***Daily Briefing***. "Precision medicine fails for up to 93% of patients. Are its proponents selling 'false hope'?". https://www.advisory.com/daily-briefing/2018/09/21/precision-medicine.

DeepMind. 2022. "AlphaFold reveals the structure of the protein universe". July 28, 2022. https://www.deepmind.com/blog/alphafold-reveals-the-structure-of-the-protein-universe.

Dorsey. 2016. State of Telehealth. N Engl J Med 2016;375:154-61. DOI: 10.1056/NEJMra1601705. http://file.medtrib.cn/pdf/61983c28-94ca-4bc6-8a5c-27191b368cdf/NEJM%20State%20of%20Telehealth%201601705.pdf .

Electronic Clinical Quality Measure (eCQM). https://ecqi.healthit.gov/glossary/cds. Reviewed February 24, 2022.

Health IT. 2016. Benefits of Electronic Health Records (EHRs). https://www.healthit.gov/providers-professionals/benefits-electronic-health-records-ehrs.

Health IT. 2016. Privacy and Security. Cybersecurity: A shared responsibility. https://www.healthit.gov/providers-professionals/cybersecurity-shared-responsibility.

Health Information and Management Systems Society (HIMSS). 2016. HIMSS Analytics 2016. ***Telemedicine Study, 2016.*** https://www.healthit.gov/sites/default/files/state_hie_evaluation_stakeholder_discussions.pdf.

HeartRhythm. 2018. Assessing the accuracy of an automated atrial fibrillation detection algorithm using smartphone technology: The iREAD Study. https://www.heartrhythmjournal.com/article/S1547-5271%2818%2930661-1/fulltext.

Hersh W. 2009. A stimulus to define informatics and health information technology. ***BMC Medical Informatics and Decision Making*** 9:24. http://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/1472-6947-9-24.

Huffington Post. 2016. The Laser Probe, The iKnife and The Cutting Edge of Surgery. ***The Huffington Post.*** http://www.huffingtonpost.com/healthline-/the-laser-probe-the-iknif_b_11803814.html.

Jie, Z., Z. Zhiying, and L. A. Li. 2021. A meta-analysis of Watson for Oncology in clinical application, ***Sci Rep***. 11(1):5792.

John Hopkins University School of Medicine (JHU). 2022. COVID-19 Dashboard. https://coronavirus.jhu.edu/map.html.

Kampov-Polevoi, J., and B. M. Hemminger. 2011. A curricula-based comparison of biomedical and health informatics programs in the USA. ***Journal of the American Medical Informatics Association*** 18(2):195–202.

Landi, H. 2022. Healthcare data breaches hit all-time high in 2021, impacting 45M people. https://www.fiercehealthcare.com/health-tech/healthcare-data-breaches-hit-all-time-high-2021-impacting-45m-people#:~:text=Cybersecurity%20breaches%20hit%20an%20all,from%2034%20million%20in%202020 .

Mesko, B. 2016a. 20 Medical Technology Advances: Medicine in the Future – Part I. http://medicalfuturist.com/20-potential-technological-advances-in-the-future-of-medicine-part-i/.

Mesko, B. 2016b. 20 Medical Technology Advances: Medicine in the Future – Part II. http://medicalfuturist.com/20-potential-technological-advances-in-the-future-of-medicine-part-ii/.

Moore, B. 2021. ArborMetrix. 2021. The Essential Guide to Modern Clinical Data Registries. Accessed August 14, 2022. https://www.arbormetrix.com/clinical-data-registries#:~:text=A%20patient%20registry%2C%20also%20called,a%20specific%20disease%20or%20condition .

National Center for Biotechnology Information (NCBI). 2016. Data and Software. https://www.ncbi.nlm.nih.gov/guide/data-software/.

National Institutes of Health (NIH). 2020. The Promise of Precision Medicine. https://www.nih.gov/about-nih/what-we-do/nih-turning-discovery-into-health/promise-precisio

n-medicine .
National Institutes of Health (NIH). 2021. Evidence Based Medicine. "Reconciling evidence and experience in the context of evidence-based practice". https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8480374/.
National Institutes of Health (NIH). 2022. List of Registries. https://www.nih.gov/health-information/nih-clinical-research-trials-you/list-registries.
National Institute of Standards and Technology (NIST). 2016. Cybersecurity framework. https://www.nist.gov/cyberframework.
National Institute of Standards and Technology (NIST). 2021. Security Measures for "EO-Critical Software" Use Under Executive Order (EO). https://www.nist.gov/system/files/documents/2021/07/09/Critical%20Software%20Use%20Security%20Measures%20Guidance.pdf .
NPJ. 2020. An overview of clinical decision support systems: benefits, risks, and strategies for success. https://www.nature.com/articles/s41746-020-0221-y.
Office of the National Coordinator for Health Information Technology (ONC). 2016. Privacy and security training games. https://www.healthit.gov/providers-professionals/privacy-security-training-games.
Office of the National Coordinator for Health Information Technology (ONC). 2020. 2020-2025 Federal Health IT Strategic Plan. https://www.healthit.gov/sites/default/files/page/2020-01/2020-2025FederalHealthIT%20StrategicPlan_0.pdf .
Oracle. n.d. "What is AI? Learn about Artificial Intelligence". https://www.oracle.com/artificial-intelligence/what-is-ai/. Accessed August 13, 2022.
Okwerekwu, J. 2016. Overlooked Warning Signs More Accurately Predict Risk of Death. ***STAT.*** https://www.statnews.com/2016/05/16/health-markers-risk-death/.
Patient Centered Primary Care Consortium (PCPCC). 2016. Defining the medical home. https://www.pcpcc.org/about/medical-home.
PC Magazine. Back Up Your Cloud: How to Download All Your Data. http://www.pcmag.com/image_popup/0,1871,s=25414&iid=408744,00.asp.
Protein Data Bank (PDB). 2022. "World Wide Protein Data Bank Archive". https://www.wwpdb.org/.
Reynolds, R. B. and M. Sharp. 2016. Health Record Content and Documentation. Chapter 4 in ***Health Information Management Concepts, Principles and Practice.*** Edited by Oachs, P.K. and Watters, A.L. Chicago, IL: AHIMA.
Swider, P. 2022. "The Future of Healthcare and Patient Care in the Metaverse". https://accelerationeconomy.com/metaverse/the-future-of-healthcare-patient.
Tory, T. 2018. ***Very Well Health***. "***Disease Management***". Overview of Disease Management. January 28, 2020. https://www.verywellhealth.com/disease-management-2615164.
University of Carolina, Davis (UC Davis). 2022. Precision medicine and its role in health care. https://health.ucdavis.edu/news/headlines/precision-medicine-and-its-role-in-health-care/2022/04.
US Department of Health and Human Services (HHS). 2010. Basic Security for the Small Healthcare Practice (Health IT Mobile Device Checklist). https://www.healthit.gov/sites/default/files/basic-security-for-the-small-healthcare-practice-checklists.pdf .
US Department of Health and Human Services (HHS). 2016. Health Resources and Services Administration. Strategic Plan 2016–2018. http://www.hrsa.gov/about/strate

gicplan/index.html.
US Department of Health and Human Services (HHS). 2020. "***Advising Providers in All Phases of Electronic Health Record Implementation***". https://www.healthit.gov/faq/what-kinds-services-do-regional-extension-centers-provide.
US Department of Health and Human Services (HHS). 2022. About Million Hearts®. https://millionhearts.hhs.gov/about-million-hearts/index.html.
US Department of Veteran Affairs (VA). 2022. VA Health Care. Updated: January 13, 2022. https://www.va.gov/health-care/.
Wyss Institute at Harvard University. 2022. Human Organs-on-Chips. https://wyss.harvard.edu/technology/human-organs-on-chips/.

ADDITIONAL RESOURCES

Bird, J. 2016. 10 Emerging healthcare trends for 2016. Healthcare. http://www.fiercehealthcare.com/healthcare/10-emerging-healthcare-trends-for-2016.
Centers for Disease Control and Prevention (CDC). 2016. Chronic diseases: The leading causes of death and disability in the United States. http://www.cdc.gov/chronicdisease/overview/index.htm.
Centers for Medicare and Medicaid Services (CMS). 2014. Eligible professional meaningful use core measures. https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/11_Clinical_Decision_Support_Rule.pdf.
Fact Sheet 19: Ethical Issues in Human Genetics and Genomics. 2016. http://www.genetics.edu.au/Publications-and-Resources/Genetics-Fact-Sheets/FactSheetELSI.
Gilpin, L. 2014. 10 Technologies changing the future of healthcare. TechRepublic. http://www.techrepublic.com/article/10-technologies-changing-the-future-of-healthcare/.
Health Resources and Services Administration (HRSA). 2016. Performance measures. http://hrsa.gov/about/strategicplan/performancemeasures.html.
IBM. 2016. ***Go beyond artificial intelligence with Watson.*** http://www/ibm.com.
IBM. 2017. A New Partnership Between Humanity and Technology. https://www.ibm.com/watson/health/.
Intel. 2016, September 21. Using wearable technology to advance Parkinson's research. http://www.healthcareitnews.com/sponsored-content/using-wearable-technology-advance-parkinsons-research.
Mell, P. 2010. Updated 2016. ***NIST SP 800-145—The NIST Definition of Cloud Computing***. National Institute of Standards and Technology. https://www.nist.gov/programs-projects/cloud-computing.
Merriam-Webster. 2010. Gamification. https://www.merriam-webster.com/dictionary/gamification.
Merriam-Webster. 2013. Genomics. http://www.merriam-webster.com/dictionary/genomics.
National Council on Aging. 2016. Chronic Disease Self-Management. https://www.ncoa.org/wp-content/uploads/Chronic-Disease-Fact-Sheet_Final-Sept-2016.pdf.
National Institutes of Health (NIH). 2015a. All about the Human Genome Project. https://www.genome.gov/10001772/all-about-the--human-genome-project-hgp/.

National Institutes of Health (NIH). 2015b. Privacy in genomics. https://www.genome.gov/27561246/privacy-in-genomics/.

National Institutes of Health (NIH). 2016a. ***Current News*** Releases. https://www.genome.gov/10000475/current-news-releases/.

National Institutes of Health (NIH). 2016b. National Human Genome Research Institute. Regulation of Genetic Tests. https://www.genome.gov/10002335/regulation-of-genetic-tests/.

National Institutes of Health (NIH). 2019. Artificial intelligence in diabetic retinopathy: A natural step to the future. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6611318/.

Office of Disease Prevention and Health Professions (ODPHP). 2016. Genomics. HealthPeople.gov. https://www.healthypeople.gov/2020/topics-objectives/topic/genomics.

US Department of Energy Office of Science. 2016. ***Human Genome Project Information*** Archive ***1990-2003, Human Genome Project.*** http://web.ornl.gov/sci/techresources/Human_Genome/index.shtml.

US Department of Health and Human Services (HHS). 2016. Regional Extension Centers. https://www.healthit.gov/providers-professionals/regional-extension-centers-recs.

Appendix A: Odd-Numbered Answers for Check Your Understanding Sections

Chapter 1

Check Your Understanding 1.1

1. d
3. d

Check Your Understanding 1.2

1. b
3. c

Check Your Understanding 1.3

1. d
3. d

Chapter 2

Check Your Understanding 2.1

1. b
3. c
5. Will vary. Students are encouraged to review and learn some of the key highlights throughout the history of EHR development.
7. d
9. e

Check Your Understanding 2.2

1. d
3. a

5.

Chapter 3

Check Your Understanding 3.1

1. a

3. d

5. a

7. Classification system provides the hierarchy of coding while terminologies codify the term within a medical information. In essence, terminologies are more granular than classifications. SNOMED CT for example contains over 300,000 terms.

Check Your Understanding 3.2

1. e

3. d

5. b

7. e

9. e

Chapter 4

Check Your Understanding 4.1

1. b

3. c

5. d

Chapter 5

Check Your Understanding 5.1

1. b

3. d

5. True

Check Your Understanding 5.2

1. b

3. a

5. d

Check Your Understanding 5.3

1. c

3. a

5. c

Chapter 6

Check Your Understanding 6.1

1. c
3. b
5. b

Check Your Understanding 6.2

1. a
3. c
5. b

Check Your Understanding 6.3

1. d
3. d

Chapter 7

Check Your Understanding 7.1

1. a
3. b
5. e
7. c

Check Your Understanding 7.2

1. False; that is the dependent variable.
3. True
5. True
7. c

Check Your Understanding 7.3

1. b
3. c
5. False
7. True

Check Your Understanding 7.4

1. a
3. d

5. a

7. d

Check Your Understanding 7.5

1. a

3. c

5. True

Check Your Understanding 7.6

1. True

Chapter 8

Check Your Understanding 8.1

1. a

3. b

5. True

7. False

Pros and Cons:

1. u

3. s

5. u

7. u

Check Your Understanding 8.2

1. d

3. c

5. a

7. b

Check Your Understanding 8.3

1. a

3. True

5. True

Chapter 9

Check Your Understanding 9.1

1. a

3. b

5. False

7. False

9. d

Check Your Understanding 9.2

1. b

3. c

5. False

7. True

9. l

11. a

13. c

Chapter 10

Check Your Understanding 10.1

1. c

3. b

Matching

5. d

7. b

9. a

Check Your Understanding 10.2

1. a

3. c

5. b

Check Your Understanding 10.3

1. c

3. b

5. c

Chapter 11

Check Your Understanding 11.1

1. c

3. b

5. d

7. True

9. False

Check Your Understanding 11.2

1. d
3. d
5. False
7. True

Chapter 12

Check Your Understanding 12.1

1. c
3. d
5. b
7. Student answers may vary. Students may choose any of the key drivers, such as cost, meaningful use, innovation, patient safety, healthcare customers, or legal/compliance pressures. Students should provide their rationale and support for their answer.

Check Your Understanding 12.2

1. b
3. a
5. b6. c

Chapter 13

Check Your Understanding 13.1

1. a
3. d

Check Your Understanding 13.2

1. d
3. d
5. b

Chapter 14

Check Your Understanding 14.1

1. e
3. d
5. c
7. b
9. American Medical Association (AMA)

11. Sample answers: Phishing; Smishing; Malware

Check Your Understanding 14.2

1. a
3. d
5. a
7. a
9. A nonprofit organization called The Clinician Taskforce
11. Ecological momentary assessment (EMA)

Chapter 15

Check Your Understanding 15.1

1. a
3. a
5. d

Check Your Understanding 15.2

1. a
3. b
5. b

Chapter 16

Check Your Understanding 16.1

1. c
3. d
5. Precision Medicine

Check Your Understanding 16.2

1. b
3. c
5. d

Appendix A: Odd-Numbered Answers for Review Exercises Sections

Chapter 1 Review Exercises

1. Health data science is a subset of health informatics that focuses on transforming health data into knowledge that can be used for making a decision.

3. In health data, there is a difference between the raw fact (data) and the meaning of that raw fact (information). For example, the raw fact contains a code of F41. 9 (data) which can mean a lot of different things, but it actually means an ICD-10 code for anxiety disorder (information).

5. New data acumen skills include the ability to make good judgments about the use of data to support evidence-based decisions, which ranges from data management, data wrangling, analytical modeling, to communications and data visualizations.

Chapter 2 Review Exercises

1. Fragmentation and decentralization of the healthcare system. 2. Licensing and accreditation processes provide insufficient attention to preventing errors. 3. Medical liability system impedes physicians openly discussing errors. 4. Overall, fewer incentives are provided for healthcare organizations and providers to improve safety.

3. Answers will vary.

5. Other settings of care utilizing EHRs include surgical centers, long-term care facilities, behavioral health therapists, dentist offices, retail pharmacy clinics, and eye-care professionals.

7. Answers will vary.

9. Answers will vary but should include these six aims of improvement:

- Safe—avoiding injuries to patients from the care that is intended to help them.
- Effective—providing services based on scientific knowledge to all who could benefit and refraining from providing services to those not likely to benefit (avoiding underuse and overuse, respectively).
- Patient-centered—providing care that is respectful of and responsive to individual patient preferences, needs, and values and ensuring that patient values guide all clinical decisions.

- Timely—reducing waits and sometimes harmful delays for both those who receive and those who give care.
- Efficient—avoiding waste, including waste of equipment, supplies, ideas, and energy.
- Equitable—providing care that does not vary in quality because of personal characteristics such as gender, ethnicity, geographic location, and socioeconomic status.

11. d

Chapter 3 Review Exercises

1. c

3. c

5. b

7. HIE was conceived prior to the HITECH legislation. Networks titled Community Health Management Information Systems (CHMISs) were first established in 1990 through grants from the Hartford Foundation. However, CHMISs experienced significant obstacles such as cost as CHMIS networks preceded the advent of widespread, reasonably priced, standards-based high-speed Internet access. In 1994, we saw the emergence of Community Health Information Networks (CHINs). These commercial initiatives focused on financial savings, with little emphasis on the public health benefits that might accrue. CHINs encountered problems when competitors limited the data sharing functionality to prevent any perceived loss of competitive advantage, the ability of vendors to charge fees, and differences in the fees charged different participants raised questions regarding any financial benefits. The late 1990s and 2000s saw the rise of the Regional Health Information Organization (RHIO) in the healthcare industry. RHIOs are the immediate precursor to the HIO. The obstacles to RHIO success were in identifying a sustainable business model that allows them to maintain the technology infrastructure, assuring privacy and security of the data, while overcoming the issue of competitor distrust. Since many RHIOs and HIOs were often created through (and operate) from a grant, these same problems exist for HIOs today when the funding dries up.

9. No consent
- Easier to administer; addresses provider concerns about incomplete information
  - Lack of patient control
- Opt-out
  - Automatic inclusion of data for more complete data for providers, yet patients can restrict use
  - Procedures for documenting opt-out so providers know about it are unclear
- Opt-out with exceptions
  - Much more patient control
  - Very, very difficult to implement
- Opt-in
  - Patient has complete, up-front control; provides an opportunity to educate patients; may provide better record matching
  - Obtaining and communicating consent; questions as to how consent can be revoked or withdrawn

- Opt-in with restrictions
  - Patient controls up front and very detailed
  - Very, very difficult to manage

11. Population health can be defined as the level and distribution of disease, functional status, and well-being of a defined group of people with specified characteristics; Public health is the study of the health of populations in large areas like states or countries.

Chapter 4 Review Exercises

1. d

3. Answer: Even though the sponsorship group approved the upgrade of the system including RF ID badges for staff, the chief human resource officer realizes that in order to maintain the quality of the implementation given budget cuts, the project implementation is not feasible including the RF ID badge software. The quality expectations stated in the case with regard to safety and security provided by pictures on staff badges will have to be delayed until the organization can afford to implement this functionality. By eliminating this requirement in the implementation plan, the CHRO provides the implementation team the opportunity to implement the software within the current stated timeframe and budget inclusive of the 10% reduction.

5. b

Chapter 5 Review Exercises

1. c

3. c

5. a

7. d

9. b

11. Data Warehous/Data marts, Boots efficiency, scalability, data security

13. Answers will vary.

Chapter 6 Review Exercises

1. d

3. a

5. b

7. a

9. a

11. Answers will vary.

13. Answers will vary.

Chapter 7 Review Exercises

1. a

3. c

5. Research question: What is the correlation between age in years and social media use per week?

Null hypothesis: There is no correlation between age in years and social media use per week.

Alternate hypothesis: There is a correlation between age in years and social media use per week.

Study variables: Age and hours of social media use per week, both categorical.

Sample answers will vary, such as population, all HIM students, or all HIM students in one university program. Data collection will vary, such as a web-based survey.

Data analysis: Pearson's correlation.

7. data mining tools include Python, IBM SPSS Data Modeler, Oracle data mining, SAS Enterprise Miner, HP Vertica Analytics Platform, and Microsoft Analysis Services. For smaller data sets, the R statistical tools can be used.

Chapter 8 Review Exercises

1. True

3. False

5. True

7. Structured and unstructured data, provide a few examples. Data are invaluable, but they typically need to be analyzed before they can provide us information.

9. A histogram can be used to show the distribution of the length of stay. A boxplot can show some statistics of the data.

11. The five characteristics of big data, one use is the data from the All of Us Research Program. Anyone can request to have access to the data and work on some projects.

Chapter 9 Review Exercises

1. b

3. c

5. False

7. False

9. True

11. False

13. False

15. Descriptive analytics is mainly used for reporting by analyzing historical data. Descriptive analytics can be used to get an aggregate-level understanding of an environment in a healthcare organization, like how many patients were admitted to the hospital, and in which department they were admitted.

Diagnostic analytics is used for diagnosing an event that has already occurred and finding the reason for the event. Diagnostic analytics will be able to find the cause of an issue based on the collected descriptive data; for instance, what is the cause of death of many patients in an intensive care unit (ICU)?

Predictive analytics uses statistical models and forecasting techniques to identify trends, correlation, and causation. Predictive analytics can be used to identify the patients who will likely be readmitted to hospitals in the next 30 days by analyzing their medical history.

Prescriptive analytics helps to identify the course of action to achieve something by leveraging optimization and simulation algorithms to advise on possible outcomes. For example, what is a more cost-effective way to manage patients at high risk for readmission: at the hospital or at the patient's home?

17. PCA projects the high-dimensional input data into lower-dimensional data by combining features. The data set thus gets smaller, with fewer important features. page 499 Just like clustering is a grouping of the data set based on distance or proximity, you could think of PCA as a partitioning of the variation in the data. It combines highly correlated variables together to form a smaller data set. In other words, PCA extracts the principal components of a data set to represent the most information in the original data.

19. a. Regression

R^2 (R-squared) is the ratio of the sum of squares regression ($SS_{Regression}$) and the sum of squares total (SS_{Total}). Sum of squares regression ($SS_{Regression}$) represents the total variation of all the predicted values found on the regression line or plane from the mean value of all the values of response variables. R^2 determines how close the data fit the regression line with a range of 0 to 1. The closer the value is to 1, the better the fit.

Here, $\hat{y}$ is the predicted value, and y is actual value of the variable. $\bar{y}$ represents the mean value of the variable.

Mean squared error (MSE) can be used to represent the cost associated with the prediction. This measures difference between the true value and the predicted value. When the linear regression model is trained using training data, the model with the lowest MSE value is selected as the best model. MSE is always positive. A perfect regression model will have zero MSE. A value closer to zero will represent better quality of the regression model.

Mean absolute error (MAE) is similar to MSE. However, MAE takes the absolute value of the sum of errors, rather than its squared sum, as in MSE. It is a more direct representation of the regression error. Whereas MAE handles all errors equally, MSE squares high prediction errors and applies a bigger penalty.

b.Classification

Accuracy

A general notion is that **accuracy** of a model can be used to evaluate the performance of a machine learning model, but accuracy can be misleading in many cases. For instance, in fraud detection or disease classification, a 90 percent accuracy may seem positive, but it might not consider the proportion of false positives or false negatives.

Accuracy can be formulated as follows:

Precision

Precision or **positive predictive value (PPV)** is another evaluation metric for classification models. Mathematically, it can be written as:

We use precision if the number of false positives is crucial.

Recall

Recall is another evaluation measure for classification models. Recall is also known as **sensitivity** or **true positive rate** and is defined as follows:

Recall is used when false negatives are critical to the application.

Specificity

Specificity is another common performance evaluation measure used for classification models. It is formulated as:

Specificity or true negative rate gives the performance of the machine learning system in predicting the negative class. In other words, when the value of specificity is high, it means that the system is good at predicting actual negative instances in the data set. In the application of detecting spam emails, high specificity means the system is good at correctly predicting spam emails correctly.

F1-Score

F1-score combines both precision and recall. This could be done by taking the mean of precision and recall.

As a harmonic mean penalizes extreme values, the F1-score uses a harmonic mean instead of averaging both values. In situations where precision is 1 and recall is 0, the harmonic mean is 0. If we were to use an average of both values, though, it would be 0.5, which is not suitable for a classifier model with high precision and low recall.

Mathematically, F1-score can be formulated as:

This can also be written as:

Area under the curve (AUC)

A receiver operating characteristic (ROC) curve can be used to evaluate the performance of a classifier. **Area under the curve (AUC)** or area under the receiver operating characteristic curve (AUROC) is the area under the ROC curve. AUC is commonly used to compare the classification performance of multiple classifiers. An AUC score of 1 is the ideal classifier.

Chapter 10 Review Exercises

1. c

3. d

5. d

7. d

9. Answers will vary.

11. Answers will vary.

13. Answers will vary.

Chapter 11 Review Exercises

1. b

3. a

5. a

7. Have you identified the ePHI within your organization? This includes ePHI that you create, receive, maintain, or transmit.

What are the external sources of ePHI? For example, do vendors or consultants create, receive, maintain, or transmit ePHI?

What are the human, natural, and environmental threats to information systems that contain ePHI?

9. Search in the news and identify one recent natural disaster. Describe the nature of the disaster, the damage to a healthcare facility, such as effects like power loss, evacuation, health care services, record access, and so on. page 501

Chapter 12 Review Exercises

1. g

3. d

5. An organization's data governance strategy should be integrated with its information governance strategy. Information governance is the specification of decision rights and an accountability framework to ensure appropriate behavior in the valuation, creation, storage, use archiving, and deletion of information, whereas data governance is the overall administration, through clearly defined procedures and plans, that assures the availability, integrity, security, and usability of the structured and unstructured data available to an organization.

7. Student answers will vary and might address any of the governance principles, such as the principles of accountability, transparency, integrity, protection, compliance, availability, retention, or disposition.

Chapter 13 Review Exercises

1. a

3. b

5. e

7. The AHIMA Standards of Ethical Coding, derived from the AHIMA Code of Ethics, sets forth the values and standards of practice for coding professionals, which serve as their guide for ethical decision-making. These standards direct coders to accurately code diagnoses and procedures and refuse to incorrectly code for increased reimbursement or other unethical reasons.

Chapter 14 Review Exercises

1. d

3. c

5. Delivering personalized communication about their healthcare, somewhat relevant to their context; Connect with their peers by sharing and learning others experience allows patients to find their community and reduce the feeling of isolation; Empower health and healthcare clients to take ownership of their own health by allowing personal health tracking and citizen-based reporting; Being able to find health-related information ready to be accessed at their fingertips; Easily complete any health-related financial transactions integrated as part of digital health solution

7. Mobile technology has a very high penetration and allow personalized intervention to be delivered just in time and customized with the patient's context in mind.

9. A patient portal is usually a type of personal health record tethered to provider's electronic health record. However, personal health record can also be a stand-alone system that is managed by the patient themselves.

Chapter 15 Review Exercises

1. Informatics is a discipline that is applicable to many different fields in healthcare. Public health informatics is an area of practice within the field of health informatics. The CDC defines public health informatics as the systematic application of knowledge about systems that capture, manage, analyze, and use information to improve population-level health outcomes. The field of public health informatics focuses on the application of informatics in public health areas, such as surveillance, prevention, preparedness, and health promotion.

3. Student answers may vary as long as students discuss any one of the 10 essential services for public health. Students should also provide rationale and support for their answer.

5. The CDC maintains several value sets in a system called the Public Health Information Network's Vocabulary Access and Distribution System (PHIN VADS). The PHIN VADS is a web-based vocabulary system that provides a standard vocabulary to the CDC and its public health partners in one place. There are currently 215 code systems, 1,944 value sets, 103 views, and more than 5 million concepts in the PHIN VADS. This system is a key component in supporting the development and deployment of standards-based public health information systems.

7. An infodemic has been defined as too much information, including false or misleading information in digital and physical environments during a disease outbreak. Infodemics lead to confusion and risky behaviors that can affect individuals' health, lead to mistrust in public authorities, and undermine the public health response. Infodemic management refers to the systematic use of risk-based and evidence-based analysis of data to manage the infodemic and reduce its impact on health behaviors during pandemics and epidemics. This is of particular concern today in the wake of the COVID-19 pandemic. The goals of infodemic management are to listen to community concerns and questions, promote understanding of risk and health expert advice, build resilience to misinformation, and engage and empower communities to take positive actions.

9. Public health surveillance has been defined as the systematic, ongoing collection, management, analysis, and interpretation of data followed by the dissemination of these data to public health programs to stimulate public health action. Surveillance of public health data such as nationally notifiable diseases is mandated at all levels to monitor, control, and prevent the spread of diseases.

Chapter 16 Review Exercises

1. b

3. c

5. NIST

7. Possible answer:

HI programs covers all topics in health informatics while certificate programs focus on specific issue or requirement in health informatics, such as Cybersecurity or Advanced data science (machine learning, AI, etc.).

9. Possible answers:

Real-time data checking, monitoring, and recording vital signs or other measures, which are important for healthcare and to determine when intervention is needed.

Artificial intelligence (AI), which is a high-level information technology used in developing machines that imitate human qualities such as learning and reasoning (ORACLE 2022).

Gamifying/gamification, which is "the process of adding games or game-like elements to something (as a task) so as to encourage participation" (2010) (Yu-Kai Chou 2022).

Holographic data input where screens and keyboards can be projected on the wall or on a table for ease and accessibility in the clinical setting.

Multifunctional radiology, which is projected to have the ability to detect many symptoms, biomarkers, and medical problems at one time and is predicted to include imaging techniques and personalized diagnostics with interventions in real-time.

In silico clinical trials that allow small microchips to be used as models of human cells, organs, or systems for simulation testing rather than the long and expensive traditional clinical trials.

Genomics and the Human Genome Project that substantiated the concept of personalized medicine with tailored therapy and tailored drugs.

Digestible sensors that function like a thin e-skin to transmit information.

Real-time diagnostics and Wearable Devices for Personal health monitoring in real-time.

MetaVerse that allow a combination of leveraging virtual reality (VR) and augmented reality (AR) technology to operate in virtual spaces.

GLOSSARY

Access:

The right to inspect and obtain a copy of protected health information about the individual in a designated record set, for as long as the protected health information is maintained in the designated record set

Accountable Care Organization (ACO):

An organization of healthcare providers accountable for the quality, cost, and overall care of Medicare beneficiaries who are assigned and enrolled in the traditional fee-for-service program

Accounting of disclosures:

A listing of all of the disclosures made outside of the entity holding the protected health information

Accreditation Standards Committee X12 (ASC X12):

Committee accredited by ANSI responsible for the development and maintenance of EDI standards for many industries. The ASC "X12N" is the subcommittee of ASC X12 responsible for the EDI health insurance administrative transactions such as 837 Institutional Health Care Claim and 835 Professional Health Care Claim forms (NIST 2011)

Addressable implementation specifications:

Standards under the HIPAA Security Rule that should be implemented unless an organization determines that the specification is not reasonable and appropriate. If this is the case, then the organization must document why it is not reasonable and appropriate and adopt an equivalent measure if it is reasonable and appropriate to do so

Ad hoc standards:

Standards established by a group of stakeholders without a formal adoption process

Administrative safeguards:

Administrative actions such as policies and procedures and documentation retention to manage the selection, development, implementation, and maintenance of security measures to protect electronic protected health information and manage the conduct of the covered entity's or business associate's workforce in relation to the protection of that information

Administrative simplification:

The category of provisions in Title II of HIPAA, in addition to the HIPAA Privacy, Security, and Enforcement Rules, Title II or the HIPAA Administrative Simplification Rule also includes rules and standards for transactions and code sets, and identifier standards for employers and providers

Aggregated health information exchange (HIE):

Combines all of the data into a centralized repository with a MPI or record locator service. The different HIE participants query, or send a request to, the repository to obtain demographic, clinical, or other information

Alert fatigue:

A commonly observed condition among physicians overwhelmed with large numbers to clinically insignificant alerts, thus causing them to "tune out" and potentially miss an important drug–drug or drug allergy alert

Alternate hypothesis:

A hypothesis that states that there is an association between independent and dependent variables

AMA Tools for Transactions and Code Set Standards:

Provides standards for the electronic exchange of protected health information, for compliance with HIPAA

Ambient intelligence:

When numerous objects operate together in harmony

American National Standards Institute (ANSI):

The US representative to the International Organization for Standardization (ISO)

American Recovery and Reinvestment Act (ARRA):

An economic stimulus bill created to help the United States recover from the downturn of the economy with a total amount of $787 billion allocated to the recovery

American Society for Testing and Materials (ASTM) on Computerized System Standards for HIT:

Provides procedures for the development, implementation, utilization, and maintenance of health information systems. Specifies data dictionary definitions to support other standards programs

ANOVA:

Test to compare two or more means to determine if they are equal or different

Any and all records:

A phrase frequently used by attorneys in the discovery phase of a legal proceeding. Subpoena-based requests containing this phrase may create a situation where the record custodian or provider's legal counsel can work to limit the records disclosed to those defined by a particular healthcare entity's legal health record. Typically, this is only during a subpoena phase, unless the information is legally privileged or similarly protected; the discovery phase of litigation probably can be used to request any and all relevant materials

Application service provider (ASP):

A third-party service company that delivers, manages, and remotely hosts standardized applications software via a network through an outsourcing contract based on fixed, monthly usage, or transaction-based pricing

Applied research:

Type of research that focuses on the use of scientific theories to improve actual practice, as in medical research applied to the treatment of patients

Architectural models:

A framework or structure for the flow of data and information within information

Archival system:

A database containing historical copies saved at a particular point in time and used later to recover and restore the information in a database system

Audit log:

Chronological record of electronic system(s) activities that enables the reconstruction, review, and examination of the sequence of events surrounding or leading to each event or transaction from its beginning to its end. Includes who performed what event and when it occurred

Authorization:

Must be obtained under the HIPAA Privacy Rule for the disclosure of PHI unless the PHI meets an exception stated in the Privacy Rule where an authorization is not required. The authorization must be written in plain language and must contain certain core elements, one of which is a statement that specifies that an authorization is signed

Autonomy:

Core ethical principle centered on the individual's right to self-determination that includes respect for the individual; in clinical applications, the patient's right to determine what does or does not happen to him or her in terms of healthcare

Availability:

Under the HIPAA Security Rule, the concept of ensuring electronic protected health information can be accessed as needed by authorized users. Encompassed in the concept of availability are many safeguards such as the data backup and storage specification

Backup:

The process of maintaining a copy of all software and data for use in the case that the primary source becomes compromised

Basic research:

A type of research that focuses on the development and refinement of theories

Beneficence:

A legal term that means promoting good for others or providing services that benefit others, such as releasing health information that will help a patient receive care or will ensure payment for services received

Big data:

"High-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making" (Gartner 2016)

Biomedical informatics:

Incorporates a core set of methodologies that are applicable for managing data, information, and knowledge across the translational medicine continuum, from bench biology to clinical care and research to public health

Blue Button:

A big, virtual button on existing patient portals in order to give patients access to their health information

Breach:

The acquisition, access, use, or disclosure of protected health information in a manner not permitted, which compromises the security or privacy of the protected health information

Break the glass:

A method that allows users to access data that they may not otherwise be allowed to access. The users are alerted in advance that their actions are being monitored giving them an opportunity to halt their actions, if inappropriate

Broad network access:

Any capabilities available over the network can be accessed by a wide variety of interface devices (laptops, smart phones, and so on) using standard mechanisms

Business associate (BA):

All entities that are not members of the workforce, but are providing a service or performing a task on behalf of a covered entity with the expectation or possibility that they will have access to protected health information

Business associate agreement (BAA):

Clearly documents the requirements of the business associate "to implement administrative, physical, and technical safeguards that reasonably and appropriately protect the confidentiality, integrity, and availability of the PHI that it creates, receives, maintains, or transmits on behalf of the covered entity" (45 CFR 164.308(b)(1) and 164.314)

Business continuity planning (BCP):

Includes the recovery and use of the technology as in disaster recovery planning as well as the ability of the organization to continue the processes required for ongoing business operations. Sometimes requires operational procedures in the absence of information technology

Business impact analysis:

A method for evaluating and prioritizing the risks that face a business. The goal of the business impact analysis is to identify any gaps in current recovery capability and to develop a strategy for meeting the identified recovery time and recovery point objectives

Cardinality:

Describes the relationship between two data tables by referring to the number of elements in each table of a database

Categorical data:

Data elements that represent mutually exclusive categories or labels

Cause and effect:

Where researcher can state with some confidence that the variable under study caused, or did not cause, the effect under study

Central Authority Platform (CAP):

Allows any or all of the classification and terminology coding systems to be used to capture and exchange clinical data

Centralized:

A system in which the systems processing functions occur on a single computer

Certification:

An evaluation performed to establish the extent to which a particular computer system, network design, or application implementation meets a prespecified set of requirements

Certification Commission for Health Information Technology (CCHIT):

An industry-wide initiative engaging a diverse group of stakeholders in a voluntary, consensus-based process that began certifying EHRs in 2006

Characteristics of data quality:

Part of the Data Quality Management Model from AHIMA. The 10 characteristics are accessibility, accuracy, consistency, comprehensiveness, currency, definition, granularity, relevancy, precision, and timeliness

Chief clinical information officer (CCIO):

Senior manager who provides nursing leadership for nurses, physicians, and other clinicians for leadership in using information systems for the delivery of care, education, and research

Chief information officer (CIO):

A senior manager responsible for the overall management of information resources in an organization

Chief knowledge officer (CKO):

The person who oversees the entire knowledge acquisition, storage, and dissemination process and identifies subject matter experts to help capture and organize the organization's knowledge assets

Chief medical information officer (CMIO):

An emerging position, typically a physician with medical informatics training that provides physician leadership and direction in the deployment of clinical applications in healthcare organizations

Chief nursing informatics officer (CNIO):

A position emerging in larger healthcare facilities and organizations. The typical responsibilities in this position are related to the implementation and utilization of HIT systems for clinical care

Civil penalties:

Generally fines or money damages used to sanction violators

Classification system:

A system that arranges or organizes like or related entities for easy retrieval

Client-server:

A network distribution method that involves multiple servers dedicated to different functions with workstations running the application and retrieving data from a server as needed

Clinical analytics:

Mining discrete patient healthcare data such as laboratory results, medication, and genetics to make clinical decisions or to aid in translating data for research or further healthcare treatment

Clinical data repository (CDR):

A central database that focuses on clinical information

Clinical data warehouse:

Snapshots of a variety of clinical databases found throughout a healthcare organization that are combined for the purpose of reporting and analysis

Clinical decision making:

The process of utilizing information to formulate a diagnosis

Clinical decision support (CDS):

The process in which individual data elements are represented in the computer by a special code to be used in making comparisons, trending results, and supplying clinical reminders and alerts

Clinical decision support system (CDSS):

A special subcategory of clinical information systems that is designed to help healthcare providers make knowledge-based clinical decisions

Clinical documentation:

The functionality of electronic capture of clinical notes

Clinical informatics:

Part of the biomedical continuum. It is focused on information systems to ensure that they support patient care that is safe, efficient, effective, timely, patient-centered, and equitable. Also, the scientific study of patient care, clinical research, and medical education and the effective use of information to support these activities, establish standards, and set policy

Clinical information systems (CIS):

Systems designed to facilitate the management of the activities of various clinical departments and to provide electronic charge capture and results reporting

Clinical outcomes assessment:

"Measures the actual outcomes of patient care and service against predetermined criteria (expected outcomes), based on the premise that care is delivered to bring about certain results." Also referred to as outcomes analysis (ONC 2012, 19)

Clinical phenotyping:

Determining which observable characteristics are applicable for a given subset of patients

Cloud computing:

A model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (such as networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction

Code set:

Any set of codes used to encode data elements, such as tables of terms, medical concepts, medical diagnostic codes, or medical procedure codes. A code set includes the codes and the descriptors of the codes

Cold site:

A type of data backup facility. Equipment must be brought in, but the site is powered and secure. This is much less expensive than other options, but could take up to a month to operationalize

Column-delimited data:

A flat file database format where information is stored in text files as columns of data. The data must be accompanied by documentation that lists the order and position of the variables so that the data may be interpreted

Comma-separated values (CSV):

A flat file database format where fields are delimited by a comma. May include the variable names in the first row. If not, then they require documentation to identify the variables in each position

Community cloud:

This cloud is supported for use by a community of users who share a trait, such as security needs, in common. One or more of the community may own and operate the cloud, or it may be operated by a third party or a combination of the two. A healthcare example of a community cloud might be the data infrastructure needed for an Accountable Care Organization or that required by a health information exchange

Comparable and consistent:

Data that has been normalized or transformed so that they conform to the designated standards

Comparative effectiveness research (CER):

Research that generates and synthesizes evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition, or to improve the delivery of care

Computer-aided diagnosis (CAD):

The incorporation of computer images with aspects of artificial intelligence to detect and identify potential disease factors

Computer-assisted coding (CAC):

Software that extracts and translates transcribed or computer-generated free-text data into diagnosis and procedural codes for billing and coding purposes

Computer-based patient record (CPR):

A historical term for an electronic patient record that provides complete and accurate data, alerts, reminders, clinical decision support, links to medical knowledge, and other aids

Computer literate:

Having sufficient knowledge and skill to be able to use computers; familiar with the operations of a computer

Computerized provider (physician) order entry (CPOE):

Applications that allow providers to write orders for medications or other treatments and transmit them electronically

Computers-on-wheels (COWs):

Self-contained rolling carts containing a computer for access to the EHR on each unit

Concept:

An idea or unique unit of knowledge or thought created by a unique combination of characteristics

Concept identification:

A text processing method that maps the text to standardized concepts using clinical terminologies

Concurrent processes:

Processes that run simultaneously and access shared resources such as databases

Confidentiality:

A legal and ethical concept that establishes the healthcare provider's responsibility for protecting health records and other personal and private information from unauthorized use or disclosure. In the context of the HIPAA Security Rule, it means that electronic health information is accessible only by authorized people and processes

Consensus standards:

Standards are those that are developed through a formal process of comment and feedback by interested stakeholders

Consolidated Health Informatics (CHI):

The initiative to adopt existing health information interoperability standards throughout all federal agencies

Consumer:

A person who purchases goods and services for personal use

Consumer informatics:

The focus is on the interest of the consumer. A goal is to support and inform consumers to facilitate the management of and participation in their own care

Continuous data:

Numerical data where there is an equal interval between the data points

Control Authority Platform (CAP):

Allows any or all of the classification and terminology coding systems to be used to capture and exchange clinical data

Control group:

A comparison study group whose members do not undergo the treatment under study

Correlational research:

A design of research that determines the existence and degree of relationships among factors

Cost of ownership:

A benefit of cloud computing. Using cloud computing can be more cost effective than maintaining a server on-site. Additional savings may be realized by reducing the number of employees and contractors needed for server maintenance, legacy system conversions, and upgrades

Covered entity (CE):

A healthcare provider, health plan, or healthcare clearinghouse that transmits health information in electronic form in connection with healthcare transactions

Criminal penalties:

A fine or imprisonment, whether suspended or not

Crosstabs:

Serve to highlight any errors where there is an expected or explicit relationship between two data elements

C-suite:

Executive management, such as the chief operating officer (COO), chief information officer (CIO), and so on

Dashboard:

Reports of process measures to help leaders follow progress to assist with strategic planning

Data:

The dates, numbers, images, symbols, letters, and words that represent basic facts and observations about people, processes, measurements, and conditions

Data analysis:

The science of examining raw data with the purpose of drawing conclusions about that information. It includes data mining, machine language, development of models, and statistical measurements. Analytics can be descriptive, predictive, or prescriptive

Data cleaning:

Sometimes termed data scrubbing, this involves examining the data thoroughly to detect wrong or inconsistent data

Data consolidation:

A process for bringing together and summarizing large amounts of data and/or data from multiple sources

Data dictionary:

A tool that provides metadata or information about data

Data flow diagram:

A method to map out a data model; it maps out the database's boundary and scope

Data governance:

Making decisions and exercising authority for data-related matters; establishing a culture where quality data is obtained and valued to drive the business

Data-interchange standards:

Establish the means by which a sender transmits or communicates data or information to a receiver (also known as transaction standards)

Data liquidity:

The ubiquitous sharing of healthcare data and information

Data mapping:

The process of associating concepts or terms from one coding system to concepts or terms in another coding system and defining their equivalence in accordance with a documented rationale and a given purpose

Data mart:

Well-organized user-centered searchable databases that contain information drawn from a data warehouse to meet the users needs. It is designed to hold the data that the department requires and is stored according to the users' specific needs

Data mining:

The process of extracting and analyzing large volumes of data from a database in order to find patterns in the data

Data model:

A representation of the data to be stored in a database and the relationships between the tables and data fields

Data Quality Management Model:

A model presented by AHIMA covering the different processes of data handling during which data quality should be addressed. The data handling processes are categorized as application, collection, warehousing, and analysis

Data retrieval:

Process of obtaining data from a healthcare database

Data segmentation:

Can be used to tag patient information so that selected portions of the record can be shared

Data set standards:

Established to assist in the tracking and understanding of important events such as births, deaths, hospital discharges, and so on. Such standards establish the data elements or data variables to be collected and define each data element

Data standards:

Standards may be related to the data set, such as the data elements specified for the Uniform Hospital Discharge Data Set, or the allowable data values, such as a selection list being provided for the address data element of State. Standards such as the allowable data elements within a data set, the definitions of the different data elements, or the allowable data values, which constrain the data that can be entered into a field, ensure that the data are as accurate as possible

Data table:

Grouping of data organized with records or rows and fields or columns

Data visualization:

The representation of data pictorially with graphs and other visual displays

Data warehouse:

A database that makes it possible to access data from multiple databases and combine the results into a single query and reporting interface

Database:

A collection of data tables

Database management system (DBMS):

A system installed locally on each machine that is designed to keep track of data locations and coordinate data modifications; creates and maintains a database. Provides a method for adding or deleting data and also supports methods to extract data for reporting

De facto standards:

Standards that have evolved over time to become universally used without a government or other mandate

Decentralized:

A system in which the processing functions are split or distributed, among one or more machines in the network system

Decision analysis:

A systematic approach to decision making under conditions of imperfect knowledge; a practical application of probability theory that is used to calculate the optimal strategy from among a series of alternative strategies

Decision support database:

A common example of a clinical data warehouse. These databases are found in many healthcare entities and may include claims data, financial data, and quality data combined in one database to support both internal and external reporting

Decision support systems (DSS):

A computer-based system that gathers data from a variety of sources and assists in providing structure to the data by using various analytical models and visual tools in order to facilitate and improve the ultimate outcome in decision-making tasks associated with nonroutine and nonrepetitive problems

Deidentified health information:

Health information that neither identifies nor provides a reasonable basis to identify an individual

Dependability:

A benefit of cloud computing. The cloud is up and running 24 hours a day, 7 days a week, every day of the year, which ensures access to much-needed items such as business applications, financial data, and help desk personnel

Dependent variable:

A measurable variable in a research study that depends on an independent variable

Descriptive research:

A type of research that determines and reports the current status of topics and subjects

Descriptive statistics:

A set of statistical techniques used to describe data such as means, frequency distributions, and standard deviations; statistical information that describes the characteristics of a specific group or a population

Designated Record Set:

A group of records maintained by or for a covered entity that is (1) the medical records and billing records about individuals maintained by or for a covered health care provider; (2) the enrollment, payment, claims adjudication, and case or medical management record systems maintained by or for a health plan; or (3) used, in whole or in part, by or for the covered entity to make decisions about individuals

Diagram 0:

A data flow diagram expanding on the context diagram and adding details regarding the data tables and their relationships

Digital Imaging Communication in Medicine (DICOM):

A standard that promotes a digital image communications format and picture archive and communications systems for use with digital images

Digital immigrant:

An individual who lived before the advent of the digital age but has adapted more or less to the digital age

Digital native:

A person who has only known a world of digital toys and tools

Disaster recovery planning (DRP):

Defines the resources, actions, tasks, and data required to manage the business recovery process in the event of a physical disaster

Disclosure:

The output and release upon request of appropriate health record documents

Discrete data:

Data elements that represent mutually exclusive categories or labels

Disease management:

Emphasizes the provider–patient relationship in the development and execution of the plan of care, prevention strategies using evidence-based guidelines to limit complications and exacerbations, and evaluation based on outcomes that support improved overall health

Distributed system:

A collection of independent computers connected through a network and managed by system software that enables the computers to coordinate their activities and to share system resources such that the users perceive the system as a single, integrated system

Document and data nonrepudiation:

Characteristics that defend against charges questioning the integrity of data or documents. They delineate the methods by which the data are maintained in an accurate form after their creation, free of unauthorized changes, modifications, updates, or similar changes

Double-blind study:

A type of clinical trial conducted with strict procedures for randomization in which neither researcher nor subject knows whether the subject is in the control group or the experimental group

e-Discovery:

The process of discovering what parts of the patient's records may be used for litigation or further legal proceedings. It is an opportunity for opposing counsel to obtain relevant information. It is named to reflect an emphasis on electronic records; however, paper–electronic hybrids may also be included

e-Health:

An emerging field at the intersection of medical informatics, public health, and business, referring to health services and information delivered or enhanced through the Internet and related technologies

e-Health literacy:

The ability to seek, find, understand, and appraise health information from electronic sources and apply the knowledge gained to addressing or solving a health problem

Electronic document management system (EDMS):

Utilized to manage the plethora of electronically stored images and documents

Electronic media:

(1) Electronic storage material on which data is or may be recorded electronically, including, for example, devices in computers (hard drives) and any removable or transportable digital memory medium, such as magnetic tape or disk, optical disk, or digital memory card; (2) transmission media used to exchange information already in electronic storage media. Transmission media include, for example, the Internet, extranet or intranet, leased lines, dial-up lines, private networks, and the physical movement of removable or transportable electronic storage media

Electronic medical record (EMR):

An electronic record of health-related information on an individual that can be created, gathered, managed, and consulted by authorized clinicians and staff within a single healthcare organization

Electronic medication administration records (eMAR):

Use of technology and bar coding to track medications from when ordered to when given to the patient

Electronic patient portals:

Provide access to personal health information from a patient's healthcare organization. These portals provide patients with opportunities to view results of exams and treatment, and also provide alerts for preventive care

Electronic protected health information (ePHI):

Individually identifiable health information that is transmitted by electronic media or maintained in electronic media

Encryption:

The process of transforming text into an unintelligible string of characters that can be transmitted via communications media with a high degree of security and then decrypted when it reaches a secure destination

Entity relationship diagram (ERD):

Displays the relationship between tables in a relational database

Episode of care:

The specific instance of a condition or illness with a defined time frame with the beginning and ending times of care identified

e-Prescribing:

Allows practitioners to use electronic devices to write and submit prescription orders directly to a participating pharmacy rather than faxing or providing the patient with a written prescription that must then be taken to the pharmacy

Equivalence:

In a data mapping, determined by the distribution of map relationships for a given map

Ergonomics:

A discipline of functional design associated with the employee in relationship to his or her work environment, including equipment, workstation, and office furniture adaptation to accommodate the employee's unique physical requirements so as to facilitate efficacy of work functions

Ethics:

A field of study that deals with moral principles, theories, and values; in healthcare, a formal decision-making process for dealing with the competing perspectives and obligations of the people who have an interest in a common problem

Evidence-based medicine:

The use of the current best evidence in making clinical decisions about the care of individual patients by integrating individual clinical expertise with the best available clinical evidence from systematic research

Experimental research:

(1) A research design used to establish cause and effect; (2) A controlled investigation in which subjects are assigned randomly to groups that experience carefully controlled interventions that are manipulated by the experimenter according to a strict protocol; ***Also called*** experimental study

Export:

To get the data out of one system

Expressivity:

How well a note conveys the patient's and provider's impressions, reasoning, and thought process; level of concern; and uncertainty to those subsequently reviewing the note

Extended or Enhanced Entity Relationship (EER):

A "languages" associated with information modeling, began as entity relationship diagramming, it is useful for those who process data better visually

External validity:

An attribute of a study's design that allows its findings to be applied to other groups

Extract, transform, and load (ETL):

Process of exporting the data from the data source, mapping the data to the data destination and importing the data into the destination data warehouse

Fault tolerant:

Systems that are highly available and remain in operation even when hardware, software, or network failures occur

Federal Rules of Civil Procedure (FRCP):

The guidelines that govern the procedures for civil trials

Federal Rules of Evidence (FRE):

Governs what and how electronic records may be used and the roles of record custodianship for purposes of evidence in a trial

Federated health information exchange (HIE):

Designed as provider-to-provider networks using the Internet for connectivity, with no central repository of data. A central entity maintains the master patient index or record locator service, which is used to determine where patients have been treated previously and where their health information might exist

Field:

Data elements representing attributes of the information being collected

Field studies:

Research done outside the laboratory or place of work

File Server:

A distribution method where the system runs entirely on the end user's workstation and transfers entire files

Firewall:

Monitors and controls all communication into and out of an intranet. It is implemented by a set of processes that act as a gateway to a network applying the organizational security policy

Flat file:

A text file, usually delimited by a comma or tab, with one record found on each row. It has only one table of data

Forecasting:

To calculate or predict some future event or condition through study and analysis of available pertinent data

Foreign key:

A variable in one table that is a primary key in another table

Forward maps:

Those that map from an older source code or data set to a newer target code or data set

Frequency:

The number of times that a particular observation or value occurs in a data set

Functional genomics:

Studies for genomics include learning the pattern of genes under certain conditions

Functionality:

An aspect of an EHR that should be evaluated. It addresses the features of the EHR product, including patient encounter documentation, automating and facilitating office workflow, decision support during patient encounters, and reporting that supports care management and template customization

Future-proofing:

A benefit of cloud computing. Computer hardware, software, and networking solutions begin their move toward obsolescence almost as soon as they are implemented, which can be costly. Cloud computing moves the cost of updating certain parts of the technology from the customer to the cloud provider, a substantial benefit given the rapid pace of the growth of the Internet and other technologies

Genetic Information Nondiscrimination Act of 2008 (GINA):

This law, which took effect in 2009, provides the legal standard to be followed for the collection, use, and disclosure of genetic information; it was enacted to prohibit discrimination based on genetic information for health coverage and employment

Genetics informatics:

Specific to the use of genome sequence data along with computer capabilities and statistical methods to obtain biological information

Genome:

All the genetic information present in a given organism

Genomics:

A branch of biotechnology concerned with the genetic mapping and DNA sequencing of sets of genes or the complete genomes of selected organisms

Government mandate:

Standards that are specified or established by the government for certain purposes

Hacker ethics:

Refers to the hackers' beliefs that restrictions on data are objectionable, and that technology is a tool that should be used to perform hands-on actions that make nonhackers' private data publicly accessible while protecting the hacker's personal privacy

Hard disk:

The permanent memory of a computer that stores installed application software, files created by the user, and the operating system

Health:

The state of being free from illness or injury

Health and Medicine Division (HMD):

(Formerly Institute of Medicine) An independent, nonprofit organization that works outside of government to provide unbiased and authoritative advice to decision makers and the public (IOM 2016)

Health informatics:

The field of information science concerned with the management of all aspects of health data and information through the application of computers and computer technologies

Health Information and Management Systems Society (HIMSS):

A cause-based not-for-profit organization exclusively focused on providing global leadership for the optimal use of IT and management systems for the betterment of healthcare. The HIMSS mission is to lead healthcare transformation through the effective use of health information technology

Health Information Exchange (HIE):

As a noun, HIE refers to an organization that supports, oversees, or governs the exchange of health-related information among organizations according to nationally recognized standards. As a verb, health information exchange refers to the actual exchange of health information electronically between providers and others with the same level of interoperability, such as labs and pharmacies

Health information system (HIS):

A set of components and procedures organized with the objective of generating information that improves healthcare management decisions at all levels of the health system

Health information technology (HIT):

The technical aspects of processing health data and records including classification and coding, abstracting, registry development, and storage

Health Information Technology for Economic and Clinical Health (HITECH) Act:

One part of ARRA. Designated funding to modernize the healthcare system by promoting and expanding the adoption of health information technology. HITECH provided $20 billion in Medicare and Medicaid incentive payments to physicians and hospitals for meaningful use of EHRs and $2.6 billion to support ONC initiatives

Health Insurance Portability and Accountability Act (HIPAA):

Legislation that addresses standards for transactions and code sets for electronic exchange of health-related information to perform billing or administrative functions; standards for terminologies related to medications including the Food and Drug Administration's names and codes for ingredients, manufactured dosage forms, drug products, and

medication packages; and additional standards for privacy and security of patient information

Health IT Standards Committee:

An official federal advisory committee established by the Health Information Technology for Economic and Clinical Health (HITECH) Act enacted as part of the American Recovery and Reinvestment Act (ARRA) of 2009. The committee makes recommendations to the National Coordinator for Health IT on standards, implementation specifications, and certification criteria for the electronic exchange and use of health information

Health Level 7 (HL7):

An international organization of healthcare professionals dedicated to creating standards for the exchange, management, and integration of electronic information

Health literacy:

The degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions

Health plan:

An individual or group plan that provides, or pays the cost of, medical care

Healthcare clearinghouse:

A public or private entity, including a billing service, repricing company, community health management information system, or community health information system, and value-added networks and switches, that do either of the following functions: (1) processes or facilitates the processing of health information received from another entity in a nonstandard format or containing nonstandard data content into standard data elements or a standard transaction; (2) receives a standard transaction from another entity and processes or facilitates the processing of health information into nonstandard format or nonstandard data content for the receiving entity

Healthcare code sets:

Any set of codes used for encoding data elements, such as tables of terms, medical concepts, medical diagnostic codes, or medical procedure codes

Healthcare Common Procedure Coding System (HCPCS):

Consists of two levels or systems. Level I is CPT, a large, well-curated catalog of clinical and surgical procedures maintained by the American Medical Association as a proprietary and copyrighted resource. Level II of the HCPCS is maintained by CMS. CMS creates and administers the Level II HCPCS codes as alphanumeric identifiers for products, supplies, and services not included in CPT

Healthcare provider:

Any person or organization who furnishes, bills, or is paid for healthcare in the normal course of business

HIPAA Enforcement Rule:

45 CFR 160, subparts C, D, and E. This rule spells out the authority of the Office for Civil Rights (OCR) related to the enforcement of the HIPAA Privacy and Security Rules

Historical research:

A research design used to investigate past events

HIT Policy Committee:

A federal advisory committee established under the HITECH Act

Hot site:

A type of data backup facility for critical applications and operations, which can be online within hours. This option is the most expensive

Human Gene Nomenclature (HUGN):

A system for exchanging information regarding the role of genes in biomedical research in the federal health sector

Hybrid cloud:

This consists of two totally separate cloud infrastructures (private, community, or public) that are unique, but share standardized or proprietary technology facilitating data and application portability

Hybrid (health) record:

A type of patient record that is maintained in both paper and electronic formats. It may be part of a mixture of paper and electronic or multiple electronic systems that do not communicate or are not logically architected for record management

Hypertext markup language (HTML):

A standardized computer language that allows the electronic transfer of information and communications among many different information systems

Hypothesis:

A statement that describes a research question in measurable terms

Imaging informatics:

Synonymous with radiology informatics or medical imaging informatics, it utilizes the radiology information gathered from individual patients to enhance patient care by finding ways to efficiently and reliably use the collected data

Import:

Brings in data from other systems or databases

Imputation:

A process that uses special statistical methods to provide missing values in data records. Imputation methods must be described fully and imputed values clearly labeled

Independent variable:

The factors in experimental research that researchers manipulate directly

Inferential statistics:

(1) Statistics that are used to make inferences from a smaller group of data to a large one; (2) A set of statistical techniques that allows researchers to make generalizations about a population's characteristics (parameters) on the basis of a sample's characteristics

Informaticist:

Individuals in a field of study (informatics) that focuses on the use of technology to improve access to, and utilization of, information

Informatics:

Using technology to acquire, manage, maintain, and use information as a basis for the plethora of decisions that must be made in a cost-effective manner. Also, the study of collecting, organizing, storing, and using electronic information

Information governance:

The accountability framework and decision rights to achieve enterprise information management (EIM). IG is the responsibility of executive leadership for developing and driving the IG strategy throughout the organization. IG encompasses both data governance (DG) and information technology governance (ITG)

Information infrastructure:

Processing, tools, and technologies to support the creation, use, transport, and storage of information

Infrastructure as a service (IaaS):

This is the most minimal level of cloud computing where the infrastructure is provided, but the consumer controls the operating systems, storage, and applications. The consumer may also have control over networking components such as the firewall

Institute of Medicine (IOM):

An independent, nonprofit organization that works outside of government to provide unbiased and authoritative advice to decision makers and the public

Institutional Review Board (IRB):

An administrative body that provides review, oversight, guidance, and approval for research projects carried out by employees serving as researchers, regardless of the location of the research (such as a university or private research agency); responsible for protecting the rights and welfare of the human subjects involved in the research. IRB oversight is mandatory for federally funded research projects

Integrated:

Designed to bring together multiple information systems and allow them to communicate in a timely and effective manner and work together as one system

Integrity:

Under the HIPAA Security Rule, this means that electronic protected health information is not altered or destroyed in an unauthorized manner

Interfaces:

Hardware or software that enable disparate CIS and HIS software systems to communicate with each other

Internal validity:

An attribute of a study's design that contributes to the accuracy of its findings

International Classification of Functioning and Disability (ICF):

Classification of health and health-related domains that describe the body functions and structures, activities, and participation

International Organization for Standardization (ISO):

A worldwide, nongovernmental, internationally recognized standards development body

Internet of Things (IoT):

A computing concept where every day physical objects are connected to the Internet and able to identify themselves to other devices

Interoperability:

Standards that allow different health information systems to work together within and across organizational boundaries in order to advance the effective delivery of healthcare for individuals and communities

Interval data:

A category of continuous data that do not have a true zero and can have negative numbers. An example of interval data is temperature, which can be negative or positive

Join:

Combine data from two or more tables in a database

Justice:

Treating all people fairly

Knowledge management (KM):

The management of information in the facility data repositories, policies and procedures, and other work documents

Laboratory information system (LIS):

A system that provides a hub to integrate laboratory information, including orders with results; there may be user access to the information. Other functionality includes scheduling, billing, and other information needed by the lab; rarely is all of this information integrated with an EHR

Legacy data:

Existing data, some of which are on paper, and some of which may already be digital. Whatever the current format, they must be converted to a format that is compatible with the new product or system

Legal health records:

A concept created to describe the data, documents, reports, and information that comprise the formal business record(s) of any healthcare organization that are to be utilized during legal proceedings

Legal hold:

A legal concept that requires special, tracked handling of patient records to ensure no changes can be made to a record involved in litigation. Common in the paper record environment to substantiate the integrity of the record, this is less common in the electronic environment where audit logs are the standard

Limited data set:

Protected health information from which certain specified direct identifiers of individuals and their relatives, household members, and employers have been removed

Linearization:

A subset of the ICD-11 foundation component, that is fit for a particular purpose (reporting mortality, morbidity, or other uses); jointly exhaustive of the ICD universe (foundation component); composed of entities that are mutually exclusive of each other; and each entity is given a single parent

Litigation response:

Typically the term used to describe the processes and procedures invoked if the potential for lawsuits or litigation is detected. It is a set of procedures that, when instituted, protects the relevant records. Litigation response is not only triggered by atypical record disclosure requests but by any sign or action that shows a potential for a lawsuit

Local area network (LAN):

A network that connects various devices together via communications within a small geographic area such as a single organization

Logical Observation Identifiers Names and Codes (LOINC):

A dictionary of laboratory codes and clinical test descriptors, originated by Regenstrief Institute and funded by federal support grants as a public resource for the common good. It is publicly usable without any fees or license

Meaningful use:

The use of a certified EHR in a meaningful manner such as for e-prescribing; clinical decision support; maintenance of a problem list of current and active diagnoses; the exchange of health information; and the submission of quality or other measures. This is evaluated as part of the EHR Incentive Program of CMS

Meaningful use of information:

Using systems with certified software to enhance the use of data obtained from health records

Measured service:

Use of resources in the cloud can be monitored, controlled, and reported, allowing for better management

Medical identity theft:

The assumption of a person's name and sometimes other parts of his or her identity—such as insurance information or Social Security number—without the victim's knowledge or consent to obtain medical services or goods, or when someone uses the person's identity to obtain money by falsifying claims for medical services and falsifying medical records to support those claims

Medical informatics:

Both the information and data parts as well as the controlling and automatic nature of data processing itself

Medical scribe:

A medical student, nurse practitioner, or physician's assistant who assists the physician with the required documentation in the patient's medical record

Metadata:

Data that provide information about another piece of data. Essentially, data about data

Method:

A piece of information stored about an object in an object-oriented database. It describes how to use the stored data

mHealth:

A form of telehealth where wireless devices and phone technologies are used for the generation, aggregation, and dissemination of health information between patients and providers

Middleware:

A bridge between two applications or the software equivalent of an interface

Minimum necessary:

Requires covered entities to evaluate their practices and enhance safeguards as needed to limit unnecessary or inappropriate access to and disclosure of protected health information

Mixed methods research:

Studies that utilize data collection and data analysis methods from both quantitative and qualitative research in one study

Mobility:

A benefit of cloud computing. Data streaming from the clouds allows for greater mobility because the Internet is accessible from multiple mobile devices, such as smart phones and tablets, PDAs, virtual desktops, or traditional laptops or PCs

Moral values:

A system of principles by which one guides one's life, usually with regard to right or wrong

Morbidity:

Diseases or conditions for which patients seek treatment

Mortality:

The cause of death on death certificates. Mortality data are used to present the characteristics of those dying, determine life expectancy, and compare mortality trends with other countries

Narrative-text string:

A type of search similar to the find functionality used every day in programs such as Microsoft Word or Adobe Reader

National Council for Prescription Drug Program (NCPDP):

A not-for-profit ANSI-accredited standards development organization founded in 1977 that develops standards for exchanging prescription and payment information (NCPDP 2017)

National Drug File–Reference Terminology (NDF-RT):

A data set created by the Veterans Administration for specific drug classifications

National Institute of Standards and Technology (NIST):

An agency of the US Department of Commerce, founded in 1901 as the nation's first federal physical science research laboratory

Natural experiments:

Where researchers have no control over the phenomenon under study. They do have the study planned but are waiting for the event to occur naturally

Natural language processing (NLP):

A technology that converts human language (structured or unstructured) into data that can be translated then manipulated by computer systems

Negation:

An issue with the widespread use of NLP in the real-time clinical domain. Negation involves the ability of the processor to detect the differences between "no chest pain," meaning it is absent, and "chest pain," meaning it is present

No-consent:

A model for HIE that does not require any agreement on the part of the patient to participate in an HIE. HIEs that have adopted this model operate in states that explicitly allow this model via legislation

Nominal data:

Those data where the categories are simply names. For example, a data element often collected for healthcare consumers is gender

Nonexperimental research:

Where there is no random selection of participants nor is there any random assignment of participants to study groups

Nonmaleficence:

A legal principle that means "first do no harm"

Notice of Privacy Practices (NPP):

An essential document regarding an individual's right to receive adequate notice of how a covered entity may use and disclose his or her PHI. The notice also must describe the individual's rights and the covered entity's legal duties with respect to that information

Null hypothesis:

A hypothesis that states there is no association between the independent and dependent variables in a research study

Numerical (computational) literacy:

The ability to calculate or reason numerically. An example of the need for this can be found in an understanding of body mass index calculations or if drug dosing may call for calculations

Nursing informatics:

Where the practice of nursing intersects with computers and information science. At the basic level it is using technology to improve patient care. The more advanced level is with the nurse scholar conducting research to validate practice and identify means of improving nursing practice to achieve enhanced outcomes and greater patient satisfaction

Object-oriented databases:

Designed to handle data types beyond text and numbers. An object-oriented database stores two types of information about the object, the data itself and the method

Object-Role Modeling (ORM):

A language associated with information modeling, sometimes called fact-oriented modeling. It uses graphical and textual languages to model and query information, as well as a procedural language for designing conceptual models

Office for Civil Rights (OCR):

The enforcement agency for the HIPAA Privacy and Security Rule. Since April 14, 2003, the Privacy Rule has been enforced by the OCR. The OCR has enforced the Security Rule since July 26, 2009

Office of the National Coordinator for Health Information Technology (ONC):

Principal adviser to the secretary of the Department of Health and Human Services on the development, application, and use of health information technology

ONC-Authorized Testing and Certification Bodies (ATCBs):

Empowered to perform complete EHR or EHR module testing and certification. They utilize conformance testing requirements, test cases, and test tools developed by the NIST to determine whether the software complies with the EHR Incentive Program requirements for establishing meaningful use of an EHR

On-demand self-service:

Anyone with the appropriate permissions can make use of the resources without additional intervention

Online diagnosers:

Individuals who use the Internet to determine a diagnosis or identify a medical condition

Ontology:

A system of categories used as a method for combining data; a common vocabulary organized by meaning that allows for an understanding of the structure of descriptive information that facilitates a specific topic or domain

Open System Interconnection Model (OSI):

Seven-layer model presenting the protocol layers that a typical data packet traverses beginning with the electronic signal that starts the source data transmission, and ending with the application layer in the destination application on a PC or server that receives the transmitted data packet and translates it into meaningful data for display (ISO 2016)

Openness:

Flexibility to extend and improve the existing system with minimal impact

Operations:

In the context of HIPAA, any of the following activities: (1) quality assessment and improvement activities, including outcomes evaluation, patient safety activities,

population-based activities, protocol development, and case management and care coordination; (2) competency assurance activities, including provider or health plan performance evaluation, credentialing, and accreditation; (3) conducting or arranging for medical reviews, audits, or legal services, including fraud and abuse detection and compliance programs; (4) specified insurance functions, such as underwriting, enrollment, premium rating, related to the creation, renewal, or replacement of health insurance or benefits, along with risk rating and reinsuring risk relating to healthcare claims; (5) business planning, development, management, and administration; and (6) business management and general administrative activities of the entity, including but not limited to deidentifying protected health information, creating a limited data set, and certain fundraising for the benefit of the covered entity

Opt-in:

A model for HIE that requires patients to specifically affirm their desire to have their data made available for exchange within an HIE. This option provides up-front control for patients since because data cannot be included unless they have agreed

Opt-in with restrictions:

A model for HIE that allows patients to make all or some defined amount of their data available for electronic exchange. Patients may restrict how their data is used by allowing access only to specific providers, by allowing only specific data elements to be included, or by allowing data to be accessed only for specific purposes

Opt-out:

A model for HIE that allows for a predetermined set of data to be automatically included in an HIE, but a patient may still deny access to information in the exchange. This model allows the consumer to prevent their data and information from inclusion in the HIE

Opt-out with exceptions:

A model for HIE that makes the patient's information available in the exchange, but enables the patient to selectively exclude data from an HIE, limit information to specific providers, or limit exchange of information to exchange for specific purposes

Ordinal data:

Data where the names or labels have an order to them with meaning attached. For example, a patient may be diagnosed with cancer, which is often represented as Stage I, Stage II, Stage III, or Stage IV

Patient-centered medical home:

Also referred to as medical home, this is a program to provide comprehensive primary care that partners physicians with the patient and their family to allow better access to healthcare and improved outcomes

Patient portal:

A secure method of communication between the healthcare provider and the patient

Payment:

Encompasses activities of a health plan to obtain premiums or to determine or fulfill its responsibility for coverage and provision of benefits under the health plan, and furnish or obtain reimbursement for healthcare delivered to an individual and activities of a healthcare provider to obtain payment or be reimbursed for the provision of healthcare to an individual

Personal health record (PHR):

An example of consumer informatics. It is an electronic or paper health record maintained and updated by an individual for himself or herself, a tool that individuals can use to

collect, track, and share past and current information about their health or the health of someone in their care

Physical models:

Capture the hardware composition of a system in terms of the computer and other devices. The model is a pictorial representation of the arrangement of the different pieces and connection points

Physical safeguards:

Physical measures, policies, and procedures to protect a covered entity's electronic information systems and related buildings and equipment, from natural and environmental hazards, and unauthorized intrusion

Physician champion:

A physician tasked with the responsibility of representing the views of the physician users

Physician Consortium for Performance Improvement (PCPI):

Develops, evaluates, uses, and disseminates best practices based on evidence-based measures

Picture archiving and communication systems (PACS):

Provides storage of electronic images and reports

Placebo:

Does not contain the study medication/intervention

Platform as a service (PaaS):

At this intermediate level of capability consumers create or acquire applications using programming languages and tools supported by the cloud computing service provider. Although the consumer does not control the underlying infrastructure they do exercise more control over the applications

Point-of-care testing (POCT):

Where the diagnostic tests are performed at or near where the patient care is taking place

Population (as related to research):

The universe of phenomena, objects, people, or data under investigation from which a sample is taken

Population health:

The capture and reporting of healthcare data that are used for public health purposes. It allows the healthcare providers to report infection diseases, immunizations, cancer, and other reportable conditions to public health officials

Post hoc text processing:

Using algorithms after the data is entered to structure the data for processing

Practicality:

An aspect of an EHR that should be evaluated. It addresses costs including goals about price, internal resources to maintain the EHR, whether the EHR is integrated or interfaced with the practice management systems, and interfaces with labs

Practice management system (PMS):

Supports scheduling, financial, and billing activities. PMSs generally include patient registration, scheduling, eligibility verification, charge capture, electronic claims processing, billing, and collections

Precision medicine:

Tailoring medical treatments to the individual characteristics of each patient through the ability to identify and classify subpopulations that differ in their susceptibility to a

specific disease, and in the biology and prognosis of conditions they may acquire, and their response to specific treatments

Predictive analysis:

Predicts future outcomes and is often used for simulating future scenarios that help managers optimize their outcomes

Predictive modeling:

Where multiple data points are analyzed utilizing sophisticated software programs to identify the one characteristic that results in a problem or provides the solution

Pre-empt:

In law, the principle that a statute at one level supersedes or is applied over the same or similar statute at a lower level

Primary data use:

Use of healthcare data for direct patient care

Primary key:

Uniquely identifies the row in a database

Privacy:

The quality or state of being hidden from, or undisturbed by, the observation or activities of other persons, or freedom from unauthorized intrusion; in healthcare-related contexts, the right of a patient to control disclosure of protected health information

Private cloud:

Accessible to only one customer and can be operated by the organization, a third party, or a combination of the two, at the customer's location or off-site. Healthcare organizations, especially larger ones, may choose this option due to the need for patient information security or if sensitive research is being conducted

Probability:

Using patient-related data to determine the likelihood of various occurrences in the disease process and modes of treatment

Protected health information (PHI):

Individually identifiable health information that is transmitted by electronic media, maintained in electronic media, or transmitted or maintained in any other form or medium

Public cloud:

This cloud is open to the general public and owned by an academic institution, the government, or some other organization for the benefit of the public. The owner or a third party or some combination thereof can operate it, at any location

Public health:

The level and distribution of disease, functional status, and well-being of all of the inhabitants of a given neighborhood, city, region, state, country, or the world

Public health informatics:

The systematic application of information and utilization of computer technology to support public health initiatives as related to research, education, and delivery of public health services

Qualitative research:

A philosophy of research that assumes that multiple contextual truths exist and bias is always present

Quality improvement (QI):

A set of activities that measures the quality of a service or product through systems or process evaluation and then implements revised processes that result in better healthcare outcomes for patients, based on standards of care

Quantitative research:

A philosophy of research that assumes that there is a single truth across time and place and that researchers are able to adopt a neutral, unbiased stance and establish causation

Quasi-experimental:

A research design that resembles experimental research but lacks random assignment to a group and manipulation of treatment

Random access memory (RAM):

Temporary memory used when a software application is in use or when the computer starts up

Random sample:

An unbiased selection of subjects that includes methods such as simple random sampling, stratified random sampling, systematic sampling, and cluster sampling

Rapid elasticity:

Service is provided on demand, meaning that more can quickly be provided when needed and reduced just as fast, so the consumer only pays for what they need and use

Ratio data:

In addition to the equal intervals between data points seen in continuous data, ratio data have a true zero (0) point. An example of ratio data in healthcare would be "patient weight"

Read only memory (ROM):

Permanent memory that can be read from but not changed by any applications' or users' actions

Record:

In database terminology, a collection of fields that are related in a database

Record custodians:

Also known as record stewards, they are commonly charged with leading multidisciplinary definition projects for both the HIPAA-required Designated Record Set and the legal health record. They are also expected to develop proper organization, management, guidelines, and processes related to the use of business records as a protective legal defense tool

Recovery point objective (RPO):

The length of time that you can operate without a particular application

Recovery time objective (RTO):

The maximum amount of time tolerable for data loss and capture

Red flags:

Suspicious documents, information, or behaviors that indicate the possibility of identity theft. Examples of red flags include inconsistencies in documents (driver's license and insurance card), documents that appear to be altered, and not having more than one form of ID (insurance card but no other form of ID)

Regional Extension Centers (RECs):

Established by the ONC to assist primary care providers in quickly becoming adept and meaningful users of EHRs. RECs provide training and support services to assist doctors

and other providers in adopting EHRs, share information and guidance to help with EHR implementation, and give technical assistance as needed

Relational database:

In such a database, data with a common purpose, concept, or source are arranged into tables. The relationship between the tables is displayed in an entity relationship diagram (ERD)

Reliability:

A measure of consistency of data items based on their reproducibility and an estimation of their error of measurement

Rendition:

The act of rendering data into documents, usually in report form. Report writers are the most common form of rendering, but there are others such as the creation of documents for continuity of care based on preprogrammed routines as required by meaningful use

Reputation:

The reputation of the vendor should also be evaluated when purchasing an EHR, including how established the vendor is, references from other clients, and confirmation that their software is certified

Request for information (RFI):

The RFI is used to ask vendors for information about their products and services. It is often used to obtain information from a large number of vendors to narrow the field of vendors to whom a request for proposals will be issued

Request for proposal (RFP):

The RFP is a formal request to vendors to provide specific information about how they can meet the organization's specific requirements. In the RFP, the organization should describe the goals and priorities for the IT acquisition including technical and functional requirements

Required implementation specifications:

Standards under the HIPAA Security Rule that are not optional, but must be implemented in conformance with the regulation

Research:

(1) An inquiry process aimed at discovering new information about a subject or revising old information. Investigation or experimentation aimed at the discovery and interpretation of facts, revision of accepted theories or laws in the light of new facts, or practical application of such new or revised theories or laws; the collecting of information about a particular subject; (2) As amended by HITECH, a systemic investigation, including research development, testing, and evaluation, designed to develop or contribute to generalized knowledge (45 CFR 164.501 2013)

Research question:

Research questions restate the research problem, name the independent and dependent variables, participants, and research site, but with a narrowed focus

Resource pooling:

The cloud computing provider resources serve multiple consumers, usually independent of location, though these can be specified by the consumer

Resource sharing:

Being able to use the hardware, software, or data anywhere in the system

Retrospective research:

A type of research conducted by reviewing records from the past (for example, birth and death certificates or health records) or by obtaining information about past events through surveys or interviews

Reverse map:

Goes from a newer source code or data set to an older target code or data set

Risk:

The probability of incurring injury or loss, the probable amount of loss foreseen by an insurer in issuing a contract, or a formal insurance term denoting liability to compensate individuals for injuries sustained in a healthcare facility

RxNORM:

A system created by the National Library of Medicine, for describing clinical drugs. RxNorm is the meaningful use specification for pharmaceutical names and codes in the United States

Sample:

A set of units selected for study that represents a population

Sanctions policy:

Under the HIPAA Security Rules, this must outline how cases of noncompliance will be addressed within the organization knowing that even with a strong risk management plan, the potential for noncompliance (either intentional or nonintentional) still exists

Scalability:

The ability to support growth while maintaining the same level of service

Scatterplot:

A visual representation of data points on an interval or ratio level used to depict relationships between two variables

Seamless care:

Consistent, coherent, logical, without discontinuities or disparities, and of uniform quality care

Secondary data use:

The nondirect care use of personal health information and other healthcare data

Secondary release of information:

When the authorized recipient will be distributing the information to other individuals or agencies

Security:

A benefit of cloud computing. The security of the cloud application is paramount for healthcare providers to be able to ensure the privacy of individually identifiable healthcare data. Cloud computing service companies working with providers and others would be subject to all aspects of the HIPAA Privacy and Security Rules

Security incident:

The attempted or successful unauthorized access, use, disclosure, modification, or destruction of information or interference with system operations in an information system

Security Risk Analysis:

Provides covered entities and business associates with the structural framework upon which to build their HIPAA Security Plan. Covered entities and business associates are required to conduct a Security Risk Analysis to evaluate risks and vulnerabilities in their environment and to implement reasonable and appropriate measures to protect against

reasonably anticipated threats or hazards to the security or integrity of electronic protected health information

Select query:

The purpose of a select query in SQL is to pull records that meet a certain criteria from a particular table or combination of tables

Semantic interoperability:

The ability to exchange data and information with its original meaning intact and understood by all receivers

Social media:

Websites and applications that enable users to create and share content or to participate in social networking

Software as a service (SaaS):

The highest level of cloud computing service models, SaaS enables access to the cloud computing provider's applications via a variety of devices, usually through an interface such as a web browser

Source:

The code or data set from which the data map originates

Spoliation:

The malicious alteration, concealment, or destruction of evidence

Standards development organizations (SDOs):

An organization involved, in the healthcare context, in creating and maintaining healthcare standards. In order to promulgate a recognized standard, SDOs must be accredited by ANSI and, if applicable, the ISO

State Rules of Civil Procedure:

State laws adopted by the state from the US Federal Rules of Civil Procedure that provide guidance regarding civil procedure within each state (US Courts 2016)

State Rules of Evidence:

State laws adapted by the state from the US Federal Rules of Civil Procedure that provide guidance regarding rules of evidence and civil procedure within each state (US Courts 2016)

Structured data:

Binary, machine-readable data in discrete fields, with limitations on what can be entered into the field. The structured entry of healthcare data involves the use of templates and on-screen forms with the possible entry fields and the potential entries in those fields controlled, defined, and limited. Structured data can have many benefits including completeness, quality, and accessibility of the data for a variety of purposes

Structured Query Language (SQL):

A fourth-generation computer language that includes both DDL and DML components and is used to create and manipulate relational databases

Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT):

A processable collection of medical terminology covering most areas of clinical information such as lab results, nonlab interventions and procedures, anatomy, diagnosis and problems, and nursing. It is a logically organized and massive terminology, mostly ordered by a formal description logic, which permits SNOMED CT to benefit from the classifying reasons developed by computer science and the fields of formal ontology and logic

Target:

The code or data set in which one is attempting to find a code or data representation with an equivalent meaning

Technical safeguards:

The technology and the policy and procedures for its use that protect electronic protected health information and control access to it

Technology literacy:

Entails having some basic knowledge about technology, some basic technical capabilities, and the ability to think critically about technological issues and act accordingly

Telehealth:

A telecommunications system that links healthcare organizations and patients from diverse geographic locations and transmits text and images for (medical) consultation and treatment

Telemedicine:

The remote delivery of healthcare services over a telecommunications network

Temporal:

An issue with the widespread use of NLP in the real-time clinical domain. Temporal means the computer can detect any time frames included when the concept is identified. For example, "past history of cancer" is very different from simply "cancer"

Terminal emulation software:

Used on a personal computer to allow a user to connect to a host as if the user is on a dumb terminal

Terminal-to-host:

The application and any data stay on a host computer with the user connecting either via a dumb terminal or using terminal emulation software on a personal computer to connect as if the user is on a dumb terminal. Sometimes this arrangement is also termed a thin client, with most of the actual processing taking place on a central computer

Terminology:

A set of terms representing the system of concepts of a particular subject field

Threat:

The potential for exploitation of a vulnerability or potential danger to a computer, network, or data

Three-tiered architecture:

Expands on the two-tier system with the addition of an application server that contains the software applications and the business rules. Clients in this architecture are called "thin clients" because most of the processing is done on the server, and the client has very little software installed on it

Topology:

The network's physical layout. Types of cabled network topologies include point-to-point, star, bus, tree, or ring

Transaction:

The transmission of information between two parties to carry out financial or administrative activities related to healthcare

Transaction set:

The data required to complete a specified communication

Transaction standards:

Establish the means by which a sender transmits or communicates data or information to a receiver (also known as data-interchange standards)

Translational medicine:

Taking the scientific knowledge gained from basic advances in research to the bedside by developing new procedures and modes of treatment to improve healthcare

Transparency:

The end user sees the distributed system as a single machine

Treatment:

The provision, coordination, or management of healthcare and related services for an individual by one or more healthcare providers, including consultation between providers regarding a patient and referral of a patient by one provider to another

T-test:

Used to compare current data to a company benchmark

Two-tiered model:

A model composed of several client computers that are used to capture and process the data. The servers in this configuration are powerful machines that have application software installed on them and in turn store the data captured by client machines

Unified Modeling Language (UML):

A common data-modeling notation used in conjunction with object-oriented database design

Uniform Rules for e-Discovery:

Amendments to the FRCP specifically designed for electronically stored information

Unintended consequences:

Unanticipated and undesirable issues or problems

Unstructured data:

Unstructured data entry is the use of free text or string data in the health record. This does not mean that there is no organization to the data, just that it is not structured with specific data elements, allowable values, and so on. In the paper record a large amount of the data, especially the day-to-day progress notes, were unstructured

Usability:

An aspect of an EHR that should be evaluated. It addresses the speed and ease of use, including goals about tasks that must be done fast, computer literacy of the practice staff, and the methods by which data will be entered

Use case:

A technique that develops scenarios based on how users will use information to assist in developing information systems that support the information requirements

Valid authorization:

Under HIPAA, a valid authorization under section 164.508 must contain at least the following elements: (i) A description of the information to be used or disclosed that identifies the information in a specific and meaningful fashion; (ii) The name or other specific identification of the person(s), or class of persons, authorized to make the requested use or disclosure; (iii) The name or other specific identification of the person(s), or the class of persons, to whom the covered entity may make the requested use or disclosure; (iv) A description of each purpose of the requested use or disclosure; (v) An expiration date or an expiration event that relates to the individual or the purpose of the use or disclose; (vi) Signature of the individual and the date. If the authorization is

signed by a personal representative of the individual, a description of such representative's authority to act for the individual must also be provided (45 CFR 164.508 2013)

Validity:

(1) The extent to which data correspond to the actual state of affairs or that an instrument measures what it purports to measure; (2) A term referring to a test's ability to accurately and consistently measure what it purports to measure

Values:

What ethics and morals are based on; the beliefs or standards, on which we base our judgment

Variable:

A characteristic or property that may take on different values

Virtual private networks (VPNs):

Extend the firewall protection boundary beyond the local intranet by the use of cryptographically protected secure channels at the IP level. A VPN is a private network that uses a public network (usually the Internet) to connect remote sites or users together privately

Visual literacy:

The ability to understand graphs or other visual information. For example, the consumer may need to understand a graph trend line of laboratory results

Vital statistics:

Statistics on events such as birth and death; important enough that a government collects information on them through certain registration procedures

Vocabulary:

A dictionary of terms; that is, words and phrases with their meanings

Vulnerability:

An inherent weakness or absence of a safeguard that could be exploited by a threat

Vulnerable subjects:

In the context of research studies are children, pregnant women, human fetuses, neonates, mentally disabled individuals, educationally or economically disadvantaged, prisoners or persons with incurable or fatal diseases. The study participants in these categories require additional consideration and/or protection

Warm site:

A type of data backup facility. IT provides basic infrastructure, but takes time, possibly a week, to activate. This is less expensive than a hot site

Whistle-blower:

One who reports a wrongdoing

Wide area network (WAN):

A computer network that connects devices across a large geographic area

Workforce:

Employees, volunteers, trainees, and other persons whose conduct, in the performance of work for a covered entity or business associate, is under the direct control of such covered entity or business associate, whether or not they are paid by the covered entity or business associate

World Wide Web (WWW):

A global network of networks offering services to users with web browsers

REFERENCES

Gartner. 2016. IT Glossary. http://blogs.gartner.com/it-glossary/big-data/.

Institute of Medicine of the National Academies. 2016. Advising the nation/Improving health. https://www.nationalacademies.org/hmd/~/media/Files/About%20the%20IOM/IOM-brochure-website.pdf.

International Standards Organization. 2016. ISO 9000- Quality Management. http://www.iso.org/iso/home/standards/management-standards/iso_9000.htm.

National Council for Prescription Drug Programs (NCPDP). 2017. About Us. http://www.ncpdp.org/About-Us.

National Institute of Standards and Technology (NIST). 2011. Accredited Standards Committee (ASC) X12. Response to the National Institute of Standards and Technology, Federal Agency Participation. https://www.nist.gov/sites/default/files/documents/standardsgov/ASC-X12-Consolidated-NIST-NFI-Response-2-16-11.pdf.

ONC. 2012. Component 5: History or Health Information Technology in the US. (Spring):19. https://www.healthit.gov/sites/default/files/Comp-05-Instructor-Manual.pdf.

US Courts. 2016. Current Rules of Practice and Procedure. http://www.uscourts.gov/rules-policies/current-rules-practice-procedure.

INDEX

A

Access controls, 309–10, 313–14, 473
Accessibility of data, 205, 348, 350, 439, 455
Accessibility of services, 460
Access to information
    in cloud computing, 458
    Cures Act on, 275–78
    denial of, 270–71
    disparities in, 149–50, 417
    by healthcare providers, 69
    in long-term care settings, 32
    management standard for, 305
    by patients, 23, 66, 146, 270–71
    right to authorize access, 271
    by unauthorized users, 301, 315
Accountability
    in access control standard, 313
    in device and media controls standard, 311
    digital health technology and, 422
    in information governance, 356
    of project managers, 88
    quality of care and, 20, 70
Accountable care organizations (ACOs), 70, 71, 114, 458, 460–62
Accounting of disclosures, 272–73, 325
Accreditation, 17, 88, 100, 145, 331, 476
Accuracy
    of data, 24, 26, 97, 145, 175, 202–3, 348, 350
    of diagnosis, 452
    of machine learning models, 247–48
ACHIPs. See AMIA Certified Health Informatics Professionals
ACOs. See Accountable care organizations
Actors, information blocking prohibitions applied to, 276–77
Acute hepatic porphyria (AHP), 244
Adaptable solutions, 107–8
Addressable implementation specifications, 299, 300, 303–7, 310–15
Address Discrepancy Rule, 322
Administrative decision support systems, 142, 145–48
Administrative functions of EHRs, 24, 26

Administrative safeguards, 260, 288, 299, 302–8
Administrative simplification, 256, 257, 266
Adoption
of CDSSs, 156–57
of classification systems, 45
of coding systems, 46, 49, 50, 52, 75
of data analytics, 251
of digital health technology, 402–3, 408, 409, 422–23
of EHRs, 3, 20–21, 28–31, 55, 166, 193, 467, 480
of health information exchanges, 57, 62–65, 67
of machine learning, 251
Rogers Adoption Curve, 92
of technology-based solutions, 90, 92–93
of tethered PHRs, 108
Advanced Encryption Standard (AES), 314, 315
Advanced Health Informatics Certification (AHIC), 476, 480
Advisory board directors, 91
Agency for Healthcare Research and Quality (AHRQ), 150, 152, 154
Agency for Toxic Substances and Disease Registry, 441–42
Aggregated HIEs, 58
Aggregation of data. See Data aggregation
AHIMA. See American Health Information Management Association
AHP. See Acute hepatic porphyria
AI. See Artificial intelligence
AIDS. See HIV/AIDS
All of Us Research Program, 225–26
AlphaFold2, 470–71
Alternate hypotheses, 174, 223
AMA. See American Medical Association
Amazon, 115
Ambulatory care settings, 27, 31–32, 70–71, 102, 185
Amendment requests, 271–72
American Health Information Management Association (AHIMA)
Body of Knowledge, 511
Career Map, 481
Certified Health Data Analyst credential, 170
Code of Ethics, 366–67, 372, 376
Consumer Health Information Bill of Rights, 274, 275
credentialing examinations offered by, 476
on data quality, 346–49
digital health as defined by, 398
Information Governance Adoption Model, 356
Information Governance Principles for Healthcare, 355–56
on legal health records, 156
MyPHR.com website, 417
RFI/RFP template, 517–37
Standards of Ethical Coding, 368
on technology-based solution adoption, 90, 92
on telemedicine, 407

American Medical Association (AMA), 50, 401–2
American Medical Informatics Association (AMIA), 368–72, 476, 480
American Recovery and Reinvestment Act of 2009 (ARRA), 6–7, 55, 256, 257, 308, 365
American Society of Health Informatics Managers, 480
American Telemedicine Association (ATA), 402
AMIA Certified Health Informatics Professionals (ACHIPs), 480
Analysis of data. See Data analysis
Analysis of variance (ANOVA) testing, 183–84
Analysis phase of SDLC, 96
Analytics. See Data analytics
Analytics and reporting verification process, 98–99
Ancillary systems, 18, 23, 28, 29
ANNs. See Artificial neural networks
Anonymized data. See Deidentified health information
Antivirus software, 365, 381, 472, 473
Apple HealthKit, 416
Application of data, in data quality management model, 346
Application programming interfaces (APIs), 56, 277–78, 407
Applied research, 167
AR. See Augmented reality
Archival systems, 457
Area under the curve (AUC), 250
ARRA. See American Recovery and Reinvestment Act of 2009
Artificial intelligence (AI). See also Natural language processing
    algorithms used in, 453, 471
    AlphaFold2 and, 470–71
    clinical decision-making and, 409
    decision support systems and, 68
    deep learning and, 180, 471
    digital health technology and, 401–2
    ethical concerns regarding, 471
    Internet of Things and, 419
    limitations of, 454
    predictive modeling and, 462
Artificial neural networks (ANNs), 185
Assessment. See also Testing
    of clinical outcomes, 171–72
    of data quality, 347
    ecological momentary, 412
    hazard vulnerability, 329
    in public health informatics, 432, 435
    of risk, 289–90, 300, 302, 320, 325–26, 329, 387
Association of Schools and Programs of Public Health, 433
Assurance, as core function of public health, 432, 433
Asynchronous interactivity, 406–9, 411, 414
ATA. See American Telemedicine Association
Attrition of subjects, as threat to internal validity, 177
AUC. See Area under the curve
Audit control standard, 314

Audit logs, 303, 374
Audits, 291–92, 304, 314, 318, 320
Audit trails, 154, 285, 381, 402, 422
Augmented reality (AR), 455–56
Authentication methods, 314, 473
Authorizations for disclosure of PHI, 167, 267–69
Automatic logout, 311, 313
Autonomy
CDSSs and, 154, 156
defined, 364
patient information and, 362, 377, 379
PHRs and, 382
in provisioning of care, 376
technology use and, 372, 403
Availability
of data, 5, 8, 44, 114, 306, 316, 327, 458
of information, 28, 32, 33
in information governance, 356
of open-source data science tools, 12

B

BAAs. See Business associate agreements
Background checks, 324, 372
Backups, 306–7, 311–12, 316, 327, 373, 457
Bar charts, 187, 216, 217, 220
BAs. See Business associates
Basic research, 167
BCP. See Business continuity planning
Behavioral healthcare, 32–33, 65, 151
Beneficence, 364, 372, 376–77, 383–84, 403, 471
BIAs. See Business impact analyses
Biases
in data, 216, 235, 237, 453, 471
of healthcare providers, 149
implicit, 6
of patients, 154
selection, 177
Big data
analysis of, 7, 68, 182, 224–26
challenges related to, 6
characteristics of, 204–5, 354
data governance of, 354
descriptive research with, 170
growth of, 11, 44, 115, 204
management of, 7
precision medicine and, 464
Bill payment services, 31, 90, 109
Biomedical informatics, 368, 432, 475–76
BI systems. See Business intelligence systems

Bivariate correlation tests, 181, 182
BLS. See Bureau of Labor Statistics
Blue Button portal, 146–48
Boards of directors, 89, 91, 95–96, 323, 326
Boxplots, 190, 222
Breaches
audits following, 291–92
as defined by HIPAA, 289, 298
notification of, 257, 290–91, 298, 319–20
risk assessment of, 289–90
types of, 298–99, 304
"Break the glass" method, 313
Bureau of Labor Statistics (BLS), 474
Business associate agreements (BAAs), 260, 308, 327
Business associates (BAs)
background checks for, 324
breaches by, 289–92, 319
as defined by HIPAA, 259, 307–8
Privacy Rule on, 257–60, 262, 273, 289–92
Security Rule on, 299–302, 307–8, 320
Business continuity planning (BCP), 328
Business impact analyses (BIAs), 326–27, 329
Business intelligence (BI) systems, 134–35

C

CAC. See Computer-assisted coding
CAHIIM. See Commission on Accreditation for Health Informatics and Information Management Education
Capital costs, 107
Cardinality, 118, 119
Care coordination
in ACOs, 460–62
aggregation of data for, 75, 458
digital health technology and, 401, 412
health information exchanges for, 62
in long-term care settings, 32, 151
medical homes and, 459–60
seamless care and, 468–69
Career Map (AHIMA), 481
Case control research, 170
Case studies
data governance vs. information governance, 357
disease prediction, 244
EHR adoption, 30–31
EHRs for research, 193
genetic predictive testing, 380
HIPAA Privacy and Security Rule violation, 318–19
hospital data quality management, 353
human resource management system upgrades, 103–4

interoperability, 74–75
litigation and metadata, 286–87
MIMIC-III data set, 227
NHANES data set, 227
normalized databases, 119–25, 503–7
patient engagement, 147–48
public health infrastructure, 438–39
quantitative research, 193
ransomware attacks, 332
real-world applications of public health informatics, 442
social networking site policies, 384–85
Stanford Medicine Research Data Repository, 136–37
telehealth, 332–33, 409–10
tethered PHR implementation, 108–9
Categorical data, 168, 174, 181, 187–88, 209–10, 213–14
Cause and effect, 170, 186
CDM. See Charge description master
CDRs. See Clinical data repositories
CDSMP. See Stanford Chronic Disease Self-Management Program
CDSSs. See Clinical decision support systems
CEHRT. See Certified Electronic Health Record Technology
Census, as data collection method, 175
Centers for Disease Control and Prevention (CDC)
applications developed by, 456
COVID Data Tracker, 442, 456
Division of Population Health, 463
EPI-Info from, 186
FluView initiative, 442
Health Literacy Action Plan, 386
ICD-10 development and, 47
National Notifiable Diseases Surveillance System, 72, 442, 539–41
on public health informatics, 432, 433
Public Health Information Network, 440
Vision for Public Health Surveillance in the 21st Century, 441–42
Centers for Medicare and Medicaid Services (CMS). See also Meaningful Use program; Medicaid; Medicare
ACOs and, 71, 461
EHR regulations from, 17, 18
FHIR use explored by, 54
Health Information Exchange Objective, 58
ICD-10 development and, 47
Innovation Center, 469
Level II HCPCS codes maintained by, 50
measures reporting, 56
patient education by, 463
Quality Assurance and Practice Improvement surveys, 143
Quality Payment Program, 114
Security Rule enforcement by, 316–17
telehealth services expanded by, 333

Value Set Authority Center and, 203
Central tendency measures, 181
CEOs. See Chief executive officers
Certification
of EHR platforms, 7, 21, 23, 28, 52, 58, 513
in health informatics fields, 476, 480
of health information technology programs, 513
Certified Electronic Health Record Technology (CEHRT), 23, 28, 58
Certified Healthcare Technology Specialists (CHTSs), 476
Certified Health Data Analysts (CHDAs), 170, 476
Certified Health Informatics Systems Professionals (CHISPs), 480
Certified Professionals in Health Informatics (CPHIs), 476
Certified Professionals in Health Information and Management Systems (CPHIMSs), 480
CEs. See Covered entities
CFOs. See Chief financial officers
Champions, of health system implementations, 87, 90, 93, 94, 423, 510
Charge description master (CDM), 117, 119, 123–24, 506–7
Charts
bar, 187, 216, 217, 220
entry misunderstandings, 66
guiding principles for, 218–19
line, 187, 220
paper, 30, 98, 166
pie, 187, 219–20
PivotCharts, 187, 191, 192
processing of, 32
CHDAs. See Certified Health Data Analysts
Checklist response format, 511–12
CHI. See Consumer health informatics
Chief executive officers (CEOs), 91, 95, 106
Chief financial officers (CFOs), 90, 95, 106–8
Chief human resource officers (CHROs), 91, 104
Chief information officers (CIOs), 96, 146, 480
Chief marketing officers, 92
Chief medical officers (CMOs), 90
Chief nursing officers, 86, 91
Chief operating officers (COOs), 146
Child mortality rates, 378
CHINs. See Community Health Information Networks
CHISPs. See Certified Health Informatics Systems Professionals
CHMISs. See Community health management information systems
Chronic disease management, 33, 149
CHROs. See Chief human resource officers
CHTSs. See Certified Healthcare Technology Specialists
CIDI. See Composite International Diagnostic Interview
CIMI. See Clinical Information Modeling Initiative
CIOs. See Chief information officers
CISs. See Clinical information systems
Civil monetary penalties (CMPs), 289, 317, 318

CLAMP tool. See Clinical Language Annotation, Modeling, and Processing tool
Classification algorithms, 242
Classification systems
adoption of, 45
disease, 247–48
in health information systems, 44
ICD, 45–49, 71, 114, 136, 145, 464, 543–44
purpose of, 45
Cleaning data, 136, 211–14, 236–37, 462
Clinical analytics, 146
Clinical data
analysis of, 30, 463
applications of, 114
collection of, 2
distribution of, 414
elements of, 45
importing, 66
models of, 53
Clinical data repositories (CDRs), 28, 29
Clinical data warehouses, 30, 134–36
Clinical decision-making
artificial intelligence and, 409
data for, 25, 61, 67, 171, 357, 364
health information exchanges and, 74
mHealth and, 413
Clinical decision support systems (CDSSs)
barriers to adoption of, 156–57
disease management and, 467–68
in EHRs, 28, 142, 151
"Five Rights" concept for, 142–43, 156
future challenges for, 157
in genomics and personalized medicine, 151–52
information access disparities for, 149–50
interoperability of, 143
in long-term care settings, 151
machine learning and, 240
patient safety and, 148–51, 157
standards for, 149
unintended consequences of, 154, 155
Clinical documentation, 18–19, 24, 29, 47, 154
Clinical informatics, 475, 476
Clinical Information Modeling Initiative (CIMI), 53
Clinical information systems (CISs), 16, 18, 20, 27, 30. See also Electronic health records
Clinical knowledge, 467, 468, 475
Clinical Laboratory Improvements Amendments of 1988, 270
Clinical Language Annotation, Modeling, and Processing (CLAMP) tool, 185
Clinical nurse specialists, 475
Clinical outcomes assessment, 171–72
Clinical Quality Framework (CQF), 149

Clinical terminologies
CPT, 50, 101, 114, 117, 136, 145
HCPCS, 50
LOINC, 51–52, 67, 152, 464
Multum, 464
overview, 49
RxNorm, 52–53, 464
SNOMED CT, 50–51, 136, 464
Clinical text mining, 185
Clinical trials, 44, 134, 136, 175, 204, 455
Clinicians. See Healthcare providers
Clinician's Guide to Use of Telehealth for CRT Service Provision (Clinician Task Force), 410
Closed-loop medication administration, 29, 30
Cloud computing
big data analytics and, 225
data availability and, 316, 327
data lakes and, 453
defined, 457–58
drawbacks of, 327–28
machine learning and, 240
NoSQL use and, 115
Clustering, 238–39, 243
CMOs. See Chief medical officers
CMPs. See Civil monetary penalties
CMS. See Centers for Medicare and Medicaid Services
CNNs. See Convolutional neural networks
Coast Guard, transfer of health data to Genesis system, 59
Codd, Edgar F., 117
Codes and coding systems. See also Healthcare code sets
adoption of, 46, 49, 50, 52, 75
computer-assisted, 145, 152–54
CPT, 50, 101, 114, 117, 136, 145
HCPCS, 50
in health information exchanges, 66–67
ICD, 45–49, 71, 114, 136, 145, 464, 543–44
LOINC, 51–52, 67, 152, 464
Multum, 464
National Drug Code, 136
proprietary, 52, 67
revenue codes, 119, 124, 503, 507–8
RxNorm, 52–53, 464
SNOMED CT, 50–51, 136, 464
standardized, 67, 75, 136
Standards of Ethical Coding, 368
Codes of ethics, 366–72, 376
Cold sites, 327
College of American Pathologists, 17, 50
Column-delimited data, 116

Comma-separated value (CSV) files, 97, 116
Commission on Accreditation for Health Informatics and Information Management Education (CAHIIM), 476
Common Rule (Federal Policy for Protection of Human Subjects), 167
Communication. See also Information and communication technology
asynchronous interactivity, 406–9, 411, 414
confidential communications requests, 274
during disasters, 330, 332
electronic, 25, 26, 56, 65, 146, 315
facilitation of, 32
with healthcare providers, 148, 274
PACS, 18, 24, 28, 29, 207, 457
with patients, 150
in public health informatics, 435
of security processes, 303
social networking and, 384
with stakeholders, 93, 100, 103
streamlining of, 32
synchronous interactivity, 405–6, 408, 411
Community-based programs, 71
Community Health Information Networks (CHINs), 57
Community health management information systems (CHMISs), 56–57, 258
Community-level disaster planning, 329
Complex rehabilitation technology (CRT), 409–10
Compliance principle for information governance, 356
Composite International Diagnostic Interview (CIDI), 151
Comprehensive care, 460
Comprehensiveness of data, 348, 350–51
Computer-assisted coding (CAC), 145, 152–54
Computerized physician order entry (CPOE), 27–29, 142, 154–55, 378, 452
Computer literacy, 416
Conceptual data models, 131
Confidential communications requests, 274
Confidentiality. See Privacy issues
Confusion matrix, 247, 248
Connected PHRs. See Tethered PHRs
Consent, 59–62, 172, 364, 384, 385
Consistency of data, 348, 351
Consolidated Clinical Data Architecture, 53
Consolidation of data, 458
Consumer health informatics (CHI), 414–15, 418
Consumer Health Information Bill of Rights (AHIMA), 274, 275
Consumers
data needs of, 347, 349–52
engagement with, 33, 66
on health information exchanges, 65–66
Context-level data flow diagrams, 131, 132
Context of information hierarchy (DIKW pyramid), 9–11
Contingency plans and operations, 306–7, 310, 325

Continuing education, 116, 304, 386, 401, 402
Continuity of care, 55, 69, 75, 321, 422, 459, 468–69
Continuous data, 174, 181, 209, 210, 212–13, 223
Control groups, 169, 177
Conversion, defined, 100. See also Data conversion
Convolutional neural networks (CNNs), 471
Cooperative HIOs, 59
Coordination of care. See Care coordination
COOs. See Chief operating officers
Coronavirus pandemic. See COVID-19 pandemic
Correlational research, 170
Correlational tests, 180–82, 221–22
Costs
    capital, 107
    of data collection, 26
    of healthcare, 7, 18, 20–21, 151–52, 256, 385, 415
    of implementation, 96, 102–3, 106–7
    operational, 107
Covered entities (CEs)
    access to PHI provided by, 270
    administrative requirements for, 287–89
    amendment requests to, 271–72
    breaches by, 289–92, 319
    business associates of, 257–60, 262, 273
    complaints filed against, 275, 288
    criminal violation enforcement and, 317
    as defined by HIPAA, 258
    disclosure of PHI by, 167, 266–69, 273
    information security programs for, 299
    policies and procedures of, 287
    reidentification mechanisms of, 263
    resolution agreements with, 289, 318
    restriction requests to, 274–75
    safeguards for, 260, 269, 288, 299, 302–15
    security risk analysis by, 300–302, 320
    workability of HIPAA regulations for, 257
COVID-19 pandemic
    adaptable solutions for, 107
    CDC tracking tool, 442, 456
    digital health technology during, 333, 403, 408–9
    health information exchanges during, 56, 63
    ICD codes related to, 47–48
    as leading cause of death, 46
    public health surveillance during, 72, 442
    quantitative research on, 168, 193
    social media use during, 384
    vaccination development, 464
    WHO dashboard on, 188
CPHIMSs. See Certified Professionals in Health Information and Management Systems

CPHIs. See Certified Professionals in Health Informatics
CPOE. See Computerized physician order entry
CPT. See Current Procedural Terminology
CQF. See Clinical Quality Framework
Criminal penalties, 317
Crosstabs, 213, 214
CRT. See Complex rehabilitation technology
C-suite, 146
CSV files. See Comma-separated value files
cTAKES project, 185
Cures Act. See 21st Century Cures Act of 2016
Currency of data, 348
Current Procedural Terminology (CPT), 50, 101, 114, 117, 136, 145
Cybersecurity, 263, 469, 471–73

D

Dashboards
  analytics, 101, 107
  Apple HealthKit, 416
  COVID-19, 188, 442
  defined, 189
  public health surveillance, 63
  for services implementation, 106
  visualization, 98, 146, 189
Data. See also Big data; Clinical data; Health data; Metadata; Structured data; Unstructured data
  accessibility of, 205, 348, 350, 439
  accuracy of, 24, 26, 97, 145, 175, 202–3, 348, 350
  availability of, 5, 8, 44, 114, 306, 316, 327, 458
  backups for, 306–7, 311–12, 316, 327, 373, 457
  biases in, 216, 235, 237, 453, 471
  capturing, 7, 11, 46, 166, 175–76, 413, 452
  categorical, 168, 174, 181, 187–88, 209–10, 213–14
  cleaning, 136, 211–14, 236–37, 462
  column-delimited, 116
  comprehensiveness of, 348, 350–51
  consistency of, 348, 351
  consolidation of, 458
  continuous, 174, 181, 209, 210, 212–13, 223
  currency of, 348
  in decision-making, 3, 5, 7, 9–12, 104, 134, 453
  defined, 10, 202, 342, 348
  demographic, 8, 149–50, 174, 209–10, 261
  in DIKW pyramid, 9–11
  discrete, 24, 29, 30, 143
  format of, 203–4
  genetic, 225, 379, 458–59
  granularity of, 348
  integration of, 11–12, 17, 151, 225, 465

interval, 174, 210
life cycle of, 3–6, 8, 203, 344
manipulation of, 98, 114, 135, 145, 166, 179, 211, 237
missing, 205, 211, 214–16, 235, 237, 458
narrative, 24
nominal, 174, 209–10, 221
numerical, 168, 210, 215, 216
operational, 114, 134
ordinal, 174, 210, 221
precision of, 348
presentation of, 192, 202, 203, 226, 452
ratio, 174, 210
raw, 114, 203–4, 211, 225, 235–37
real-time, 354, 454
redundancy of, 73, 117, 118, 458
relevance of, 349
reporting, 44–45, 56, 72–73, 191–93, 204
retrieval of, 189–90
segmentation of, 65
standards for, 44, 176, 202–3, 350, 433, 440
storage of, 311, 312, 316, 318–19, 457–58
timeliness of, 349
transformation into knowledge, 2, 5, 9–12, 95
value of, 205, 354
variety of, 179, 205, 224, 354
velocity of, 204, 205, 354
veracity of, 205, 354
visualization of, 180, 187–89, 192, 216–20, 226
volume of, 7, 11, 193, 205, 225, 354, 474–75
Data acumen, 12
Data aggregation
for care coordination, 75, 458
from digital health technology, 403
interoperability and, 69
for mortality and morbidity statistics, 45
for population health decision-making, 463–64
privacy and security issues in, 71
in SmartCare concept, 419
for value-based care initiatives, 70
Data analysis
advanced, 182–86
of big data, 7, 68, 182, 224–26
cleaning data for, 211–14
of clinical data, 30, 463
in data quality management model, 346
defined, 176
exploratory, 220–22, 227
forecasting in, 184
imputation and, 215–16, 237

machine learning in, 5, 11–12, 227
predictive, 184
preprocessing procedure, 211, 236–37
in research, 168–70, 176–86
results presentation, 226
retrospective, 47
statistical, 176–86, 191, 215
trend, 183–84
understanding data for, 211–12
Data analytics, 234–39
challenges related to, 251
for decision-making, 453
defined, 146, 234
for disease prediction, 244
in healthcare, 250–51
in public health informatics, 433, 435
steps for, 235–39
training in, 98, 108
types of, 234–35
Database administrators (DBAs), 128, 191–93
Database management software, 120, 133
Database management systems (DBMSs), 5, 114–15, 126, 128, 132–33, 191
Databases, 113–37. See also Relational databases; specific databases
backups for, 457
case studies, 119–25, 136–37, 503–7
data warehouses and marts, 134–36
decision support, 135
flat file, 115–17, 119, 120, 503
manipulation of, 115, 120, 126
modeling process for, 117, 131–33
normalized, 118–25, 237, 353, 464, 503–7
NoSQL, 5, 115
object-oriented, 5, 120
retrieval of data from, 189–90
for secondary use, 44
SQL, 126, 128–30, 190
terminology for, 114–15
types of, 115–20
Data collection
in analysis phase of SDLC, 96
burden of, 26
in data analytics, 235
in data quality management model, 346
instruments for, 175–76
from medical devices, 24, 205
methods for, 168, 169, 175, 203, 205, 212
from patient portals, 463
on post-conversion support, 101
in public health informatics, 439

purposes of, 44, 202, 204
in real-time, 454
research design and, 168–71
of structured vs. unstructured data, 205–7
Data conversion, 97–98, 511, 515
Data dictionaries, 116, 126–27, 131, 192, 211–13, 221
Data entry, 24, 50, 66, 205, 207
Data exchange. See also Health information exchanges
consumer-mediated, 66
decision support systems and, 53
importance of, 4
LOINC and, 52
in long-term care settings, 32
of medical records, 53
RxNorm and, 52
Data flow diagrams, 131, 132
Data governance, 341–57
of big data, 354
comparison with data management, 342
consumer data needs and, 347, 349–52
data marts as challenge for, 135
definitions of, 342
framework for, 344, 345
functions of, 342–43
information governance vs., 354–57
management of, 353–54
process for, 343–44
quality of data and, 346–51
structure for, 354, 355
Data handling processes, 346
Data integrity
copy/paste function and, 377
database design and, 117, 118
data marts and, 193
facilities reports and, 192
Security Rule on, 314–16
of structured data, 342
Data lakes, 235, 453, 458
Data level security, 99
Data literacy, 12
Data management
advancements in, 316, 463
of big data, 7
comparison with data governance, 342
data dictionaries and, 126
defined, 342
drivers of effectiveness, 343
improper, 237
Data mapping, 135, 220

Data marts, 135, 193, 353
Data mining, 143, 146, 151–54, 184–86, 220, 454
Data models, 53, 106, 117, 131–33, 344, 463
Data quality
assessment of, 347
in big data, 205, 225
in CDSSs, 156
characteristics of, 346–51
collection methods and, 203
in data governance framework, 344
defined, 346
digital health technology and, 401
for implementation of health systems, 98
interoperability and, 71
management of, 346–47, 353
preprocessing and, 227
Data science
DIKW pyramid and, 10–11
health, 2, 3, 9, 11–12
health informatics in relation to, 2, 3
interdisciplinary nature of, 202
open-source tools for, 12
Data scientists, 179, 189, 398, 400, 401, 423
Data tables, 114–26, 226, 503–7
Data Use and Reciprocal Support Agreement (DURSA), 62
Data warehouses
administrative decision support systems and, 146
centralized, 71
clinical, 30, 134–36
in data quality management model, 346
defined, 235
updating, 353
DBAs. See Database administrators
DBMSs. See Database management systems
Death rates. See Mortality rates
Decision-making. See also Clinical decision-making
data-driven, 3, 5, 7, 9–12, 104, 134, 453
ethical, 363, 365, 374–76, 379, 387–89
evidence-based, 12, 90, 202
governance structures for, 88, 90
implicit bias as influence on, 6
informed, 16, 400
organizational, 88–90, 115, 142
for population health, 463–64
structured data in, 146
transparency in, 374
Decision support databases, 135
Decision support systems (DSSs), 141–57. See also Clinical decision support systems
administrative, 142, 145–48

in behavioral healthcare, 151
in computer-assisted coding, 145, 152–54
data exchange and, 53
in EHRs, 25, 26, 28, 31, 97, 142
knowledge management and, 143–44, 154
machine learning and, 68, 240
patient safety and, 142, 146, 148–51, 157
quality of care and, 142–46, 148, 155–56
research on use of, 150–54
Decision trees, 185, 466
Decryption, 313–15
Deep learning, 7, 180, 207, 471
DeepMind, 470–71
Defense Department, 59
Deidentified health information, 135–36, 167, 172, 227, 263–66
Delaware Health Information Network, 61, 67
Demographic data, 8, 149–50, 174, 209–10, 261
Density estimation, 243
Department of ::. See specific name of department
Dependent variables, 173, 174, 180, 183, 223–24
Depersonalization, 409
Descriptive analytics, 234
Descriptive research, 170
Descriptive statistics, 176, 181, 212–13, 220–22
Designated record sets (DRSs), 153, 270–72, 276
Device and media controls standard, 311–12
Diagnostic analytics, 234
Diagnostic coding system. See ICD-10-CM
Diagram 0 data flow diagrams, 132
DICOM. See Digital Imaging and Communications in Medicine
Differential selection of research participants, 177
Diffusion of treatment, as threat to internal validity, 177
Digestible sensors, 455
Digital health literacy, 423
Digital health technology, 397–423. See also Electronic health records; Medical devices; Telehealth; Telemedicine; Wearable devices
adoption of, 402–3, 408, 409, 422–23
asynchronous interactivity in, 406–9, 411, 414
challenges related to, 421–22
characteristics of, 399
consumer health informatics and, 414–15, 418
in COVID-19 pandemic, 333, 403, 408–9
defined, 398
diagnostic tools, 453
e-health, 399, 413–14
EMRs, 18–20, 28–31, 114, 376, 417
ethics of, 403
health apps, 33, 225, 411, 413
integration of, 399–401, 418–19

mHealth, 401–2, 410–13, 419
origins and evolution of, 398–99
PHRs and, 415–18
professional organizations on, 401–2
stakeholders in, 398, 400–401, 421–23
synchronous interactivity in, 405–6, 408, 411
telerehabilitation, 399, 403, 406
videoconferencing, 406, 408, 410
Digital Imaging and Communications in Medicine (DICOM), 207
DIKW pyramid (context of information hierarchy), 9–11
Disaster planning, 325–32
community-level, 329
health information exchanges and, 59, 73
in implementation phase of SDLC, 99
for information asset protection, 325–28
phases of, 329–32
Disaster recovery planning (DRP), 306–7, 310, 328, 331–32
Disclosure of protected health information
accounting of disclosures, 272–73, 325
authorization for, 167, 267
as defined by HIPAA, 267, 272
incidental use and, 268–69, 273
limited data sets, 269, 273
mitigation for wrongful disclosures, 288
opportunity to agree or object to, 268
for public interest purposes, 269
restriction requests, 274–75
to subject of PHI, 267, 273
for TPO, 268, 272–74
Discrete data, 24, 29, 30, 143
Disease classification, 247–48
Disease management (DM), 31, 33, 149, 202, 204, 466–68
Disease prediction, 244
Disease registries, 441, 466–67
Disposal implementation specification, 312
Disposition principle for information governance, 356
Doctors. See Healthcare providers
Documentation. See also Records
of access controls, 310
of accounting of disclosures, 273
clinical, 18–19, 24, 29, 47, 154
for column-delimited data, 116
of consumer data needs, 351
copy/paste shortcut for, 377
of episodes of care, 16
errors in, 154, 377
in health records, 148
of HIPAA training, 288
of implementations, 100, 108

MU program on, 20
online systems for, 263
of security incidents, 306
of security risk analysis, 302
of security training, 305
template-driven initiatives, 18
of test irregularities, 98
of vendor selection, 513
Double-blind studies, 169
Downtime and disaster planning verification, 99
DRP. See Disaster recovery planning
DRSs. See Designated record sets
Drugs. See Medications
DSSs. See Decision support systems
Dunnett's C test, 184
DURSA. See Data Use and Reciprocal Support Agreement

E

Early adopters, 92–93
EBM. See Evidence-based medicine
Ebola virus, 71–72
Ecological momentary assessment (EMA), 412
eCQM. See Electronic clinical quality measures
EDA. See Exploratory data analysis
E-discovery, 285–86, 316
EDISs. See Emergency department information systems
EDs. See Emergency departments
Education and training. See also Patient education
on coding consistency, 67
continuing, 116, 304, 386, 401, 402
for disaster preparedness, 330
on ethical issues, 388
in health informatics, 453, 475–76, 483
for health literacy, 386
for health system implementation, 87, 96, 99–100
on HIPAA policies, 288, 318, 373
on medical identity theft prevention, 323, 324
on patient portal use, 146
on public health initiatives, 46, 73
in reporting and data analytics, 98, 108
on security, 305–6
for superusers, 108
Efficiency of healthcare
decision support systems and, 146
digital health technology and, 401, 422
health information exchanges and, 64
improvements in, 114, 250, 459
interoperability and, 69
in long-term care settings, 32

technology platforms and, 17, 20–22, 25, 26
tethered PHRs and, 33
Efficiency of technology adoption, 92, 93
E-health, 399, 413–14
eHealth Exchange, 63
EHR System Functional Model (EHR-S FM), 512
Electronic clinical quality measures (eCQM), 149
Electronic communications, 25, 26, 56, 65, 146, 315
Electronic health information (EHI), 23, 66, 143, 262, 266, 275–78
Electronic health records (EHRs), 17–36
adoption of, 3, 20–21, 28–31, 55, 166, 193, 467, 480
in ambulatory care settings, 27, 31–32, 102
in behavioral healthcare, 32–33, 151
case study, 30–31
certification of platforms, 7, 21, 23, 28, 52, 58, 513
clinical functions of, 25–26
components of, 23–24
consolidation of, 458
consumer perceptions of, 66
copying/pasting data to and from EHRs, 377
data availability from, 114
decision support systems in, 25, 26, 28, 31, 97, 142
defined, 5, 17, 166, 381
for disease management, 31, 202
ethical concerns regarding, 381–82
functional requirement specifications, 512
HHS on measuring use of, 467
history of, 18–20
implementation of, 87, 88, 96, 102
in inpatient care settings, 27–30
integration of systems, 21, 23, 24, 30–31
interoperability of, 17, 20–21, 23, 52–55
in long-term care settings, 32–33
MU program and, 7, 20–21, 28, 55, 382
order sets within, 97, 107
primary uses of, 25, 166
privacy of, 7, 23, 24, 288
qualitative research on, 21–22
for research, 193
RFI/RFP process for, 509–15
secondary uses of, 25, 166
security issues involving, 23, 24
sentinel events involving, 378–79
updates and upgrades to, 105
vendors of, 21, 23, 31, 148
Electronic media, 257, 261–63, 306, 311–12, 318, 319
Electronic medical records (EMRs), 18–20, 28–31, 114, 376, 417
Electronic medication administration records (eMARs), 28, 29
Electronic patient portals. See Patient portals

Electronic prescribing. See E-prescribing
Electronic protected health information (ePHI)
breaches related to, 298–99
as defined by HIPAA, 262–63
hacker ethics applied to, 381
ransomware attacks and, 332
Security Rule on, 262–63, 299–315
EMA. See Ecological momentary assessment
eMARs. See Electronic medication administration records
Emergency access procedures, 313
Emergency department information systems (EDISs), 19
Emergency departments (EDs)
disaster planning and, 329
EHR solutions in support of, 17, 30
health information exchange use in, 63
public health data reporting from, 73
tracking patient usage of, 59, 63
value-based care initiatives for, 70
Emergency mode operation plans, 306–7, 310
Emerging technology, 453–59, 469–73
Employees. See Workforce
EMR Adoption Model (EMRAM[SM]), 28–31
EMRs. See Electronic medical records
Encryption, 290, 313–16, 319, 332, 374, 381, 458, 473
End Stage Renal Disease Seamless Care Organizations (ESCOs), 469
Enforcement Rule (HIPAA), 289, 299, 317
Entity authentication standard, 314
Entity relationship diagrams (ERDs), 117, 118, 120, 125, 131, 132
Environmental threats to security, 301
EOBs. See Explanations of Benefits
ePHI. See Electronic protected health information
Epidemiological research, 186
EPI-Info, 186
Episodes of care, 16, 18, 166, 400
EpiTools package, 186
E-prescribing, 20, 30–32, 52, 381
Errors
in data, 211–14, 236–37, 463
in documentation, 154, 377
medical, 21, 346, 378, 381, 460
in query codes, 192–93
ESCOs. See End Stage Renal Disease Seamless Care Organizations
Ethics, 361–89. See also Consent; Privacy issues
artificial intelligence and, 471
codes of, 366–72, 376
in decision-making, 363, 365, 374–76, 379, 387–89
defining, 362–63
of digital health technology, 403
hacker, 381, 472

in healthcare, 375–86
legal influences on, 365–66
machine learning and, 471
medical, 257, 362, 403
policies and procedures for, 372–74
principles related to, 364–65
in standards of practice, 366–74
Ethics committees, 374–75
ETL process. See Extract, transform, and load process
Evaluation standard, 307
Evidence-based decision-making, 12, 90, 202
Evidence-based medicine (EBM), 29, 455, 460, 468
Experimental research, 169–70
Explanations of Benefits (EOBs), 54, 324, 383
Exploratory data analysis (EDA), 220–22, 227
Export function, 190
External stakeholders, 87
External validity, 177
Extract, transform, and load (ETL) process, 135

F

Facebook, 65, 146, 384
Facility access control standard, 309–10
Facility security plans, 310
Fair Credit Reporting Act of 1970, 322
False negatives (FNs), 247–49
False positives (FPs), 247–49
Family Educational Rights and Privacy Act of 1974 (FERPA), 262, 276
Fast Healthcare Interoperability Resources (FHIR), 54, 143, 176, 278
FDA. See Food and Drug Administration
FDA Adverse Event Reporting System (FAERS), 53
Feature engineering, 237–39
Federal Deposit Insurance Corporation, 322
Federal Policy for Protection of Human Subjects (Common Rule), 167
Federal Reserve System, 322
Federal Rules of Civil Procedure, 285
Federal Trade Commission (FTC), 291, 322
Federated HIEs, 58
Fee-for-service programs, 21
FERPA. See Family Educational Rights and Privacy Act of 1974
FHIR. See Fast Healthcare Interoperability Resources
Fields, in data tables, 114–17, 126
Field studies, 171
First Databank, 53
First normalized form, 118, 119, 503
"Five Rights" concept, 29, 142–43, 156
Five Vs of Data, 354
Flat file databases, 115–17, 119, 120, 503
FluView initiative, 442

FNs. See False negatives
Focus groups, 168, 169
F1-score, 249
Food and Drug Administration (FDA), 53, 269, 413
Forecasting, 184
Foreign keys, 117–19, 126, 191, 503
Foundation Layer of ICD-11, 48, 49
FPs. See False positives
Freedom of choice. See Autonomy
Frequency of data elements, 187, 212–15, 221
FTC. See Federal Trade Commission
Functional Profile Registry, 512

G

Gamifying, defined, 454
General Equivalency Mappings (GEMs), 47
Genesis system, 59
Genetic data, 225, 379, 458–59
Genetic information, 257, 286, 308, 366, 379
Genetic Information Nondiscrimination Act of 2008 (GINA), 257, 286, 308, 366
Genetic testing, 286, 366, 379–80
Genomics
    data, 8, 49, 205, 225, 458–59, 465–66
    personalized medicine and, 151–52, 455
    pharmacogenomics, 225, 464–65
Geospatial maps, 188
Globalization, 376, 442, 444–46
Google, 115
Governance
    of data. See Data governance
    of information, 135, 354–57
    organizational, 88–90, 97, 99, 106–8
Grand Rounds, 456
Granularity of data, 348
Graphs, 183, 187–88, 191, 192, 218–20

H

Hackers, 299, 301, 325–26, 332, 381, 402, 472
Hall, Edward, 412
Hardware
    for database housing, 132, 133
    device and media controls standard for, 311
    implementation of, 105–6
Hash function, 315
Hassabis, Demis, 471
Hazard vulnerability assessment (HVA), 329
HCPCS. See Healthcare Common Procedure Coding System
HD. See Huntington's disease
Health. See also Population health; Public health; Telehealth

e-health, 399, 413–14
mHealth, 401–2, 410–13, 419
social determinants of, 32–33, 239, 444
Health and Human Services Department (HHS). See also Meaningful Use program; Office of the National Coordinator for Health Information Technology
Common Rule and, 167
on deidentified health information, 263
health information exchanges and, 152
health literacy action plan, 385–86
Healthy People 2030, 438–39
on measuring use of EHRs, 467
Office for Civil Rights, 257, 263, 275, 289, 291, 298, 316–18
Office for Human Research Protections, 171–72
Office of Disease Prevention and Health Promotion, 433, 438
Office of Minority Health, 150
Office of the Inspector General, 278
resolution agreements with, 318
Healthcare. See also Healthcare providers; Medicine
behavioral, 32–33, 65, 151
continuity of, 55, 69, 75, 321, 422, 459, 468–69
coordination of. See Care coordination
costs of, 7, 18, 20–21, 151–52, 256, 385, 415
data analytics in, 250–51
digitization of, 2, 3, 6–7, 11, 20, 408
disparities in, 149–50, 414, 417, 461
efficiency of. See Efficiency of healthcare
emerging technology in, 453–59, 469–73
episodes of care, 16, 18, 166, 400
ethical issues in, 375–86
forecasting in, 184
incentives in payment models for, 63–64
machine learning within, 240
patient-centered, 20, 386, 467, 469
preventive, 71, 142, 244, 464–66
provisioning of, 71, 142, 364, 375, 376, 461
quality of. See Quality of care
transactions in, 5, 258, 265, 266
trends in, 459–69
Healthcare clearinghouses, 167, 257, 258, 305
Healthcare code sets. See also Codes and coding systems
defined, 44–45
ethical standards and, 368
HIPAA rule on, 298
mapping between, 136
standards for, 256
Healthcare Common Procedure Coding System (HCPCS), 50
Healthcare operations, 268, 269, 272, 273, 364
Healthcare organizations
accreditation of, 100, 145, 331

adaptability of, 107
disaster planning for, 328–32
diversity of, 104
EMR adoption by, 28
patient engagement by, 146–48, 150, 151, 154
population health initiatives of, 70
Red Flag Rules for, 322–23
unstructured data within, 8
Healthcare providers. See also Healthcare; Medicine
access to health information, 69
in ambulatory care settings, 31, 71
biases of, 149
communication with, 148, 274
compliance with HIPAA Privacy Rule, 257
cooperative HIOs among, 59
as defined by HIPAA, 258
digital health needs of, 400–401
on health information exchanges, 64–65
monetary incentives for EHR use, 7
operational workflow of, 21, 63
public health data reporting by, 73
reimbursement from Medicare, 23
value-based care initiatives of, 70
Healthcare transactions, 5, 258, 265, 266
Health data. See also Clinical data
capturing, 7, 11, 46, 166, 175–76, 413, 452
challenges unique to, 7–8
generation of, 4, 6–7, 418
genomic, 8, 49, 205, 225, 458–59, 465–66
interoperability of, 59
laboratory, 49, 51–52, 73, 212, 439
life cycle of, 3–6, 8
medical imaging, 7, 29–30, 207, 457
patient-generated, 383, 402, 414, 418
primary uses of, 25, 166, 408
public health, 44–46, 56, 72–73, 433, 439, 440
for reimbursement, 4, 43, 377
secondary uses of, 25, 44, 50, 166, 204, 206, 263
sensitivity of, 65
Health data science, 2, 3, 9, 11–12
Healtheway, Inc., 62–63
HealthGrades website, 146
Health informatics (HI). See also Public health informatics
benefits of, 452–53
certifications within, 476, 480
consumer, 414–15, 418
core skills of, 11–12
data science in relation to, 2, 3
defined, 2, 16, 86

DIKW pyramid in, 9–11
disaster preparedness and, 325
education in, 453, 475–76, 483
forces shaping field of, 6–7
functions of, 321
medical identity theft and, 324–25
origins and evolution of, 2, 3
population health, 443–44
practitioners of, 71
statistical analysis software used in, 178
workforce in, 474–81
Health information
access to, 23, 69, 149–50, 417
as clinical function of EHRs, 25, 26
deidentified, 135–36, 167, 172, 227, 263–66
electronic, 23, 66, 143, 262, 266, 275–78
ethical use of, 376–79
privacy of. See Privacy issues
protected. See Protected health information
public health benefits of, 57
security of. See Security issues
sensitivity of, 65
Health Information and Management Systems Society (HIMSS), 28–31, 480
Health information exchange organizations (HIOs), 60–62
defined, 44, 55
financial sustainability of, 67
future challenges for, 66–68, 205
organizational structure of, 58–59
patient consent options for, 59–62
public health benefits of, 72–74
RHIOs as precursor to, 57
stakeholder perceptions of, 64–65
Health information exchanges (HIEs), 55–68
adoption of, 57, 62–65, 67
consumer perceptions of, 65–66
for continuity of care, 75
Cures Act on, 45, 275
digital health technology and, 399
ethical concerns regarding, 383–84
framework for development of, 54
future challenges for, 66–68
guidelines for continued development of, 57
in HIMSS EMR Adoption Model, 29, 30
interoperability and, 55, 56, 59
MU program and, 55–57, 63
origins and evolution of, 56–58
patient consent to, 59–62, 384
privacy of, 57, 58
public health and, 58, 59, 72–74, 440

RECs and, 467
return on investment in, 67
security of, 57, 58, 152
stakeholder perceptions of, 64–65
technical structure of, 58–59
Health information systems. See also Implementation of health information systems
community-based, 56–57, 258
consumer interaction with, 414
electronic, 452–53
encryption of, 314
functions of, 2
privacy issues in, 5
public, 433, 440
security issues in, 5, 314
standards used in, 44
Health information technology (HIT)
advancements in, 417
big data resulting from, 44, 204
certification programs for, 513
implementation of, 64
information access disparities and, 149, 150, 417
infrastructure for, 45
sentinel events and, 377–78
Health Information Technology Advisory Committee (HITAC), 45
Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009
on accounting of disclosures, 273
on breach notification, 291, 298, 319
on business associate agreements, 260
healthcare transactions as defined by, 266
on health information exchanges, 55, 57, 62, 67
on health information technology certification, 513
HIPAA enforcement rules strengthened by, 257
Meaningful Use program and, 20
overview, 365–66
REC funding from, 467
research as defined by, 166
on segmentation of data, 65
Health Information Technology Policy Committee (HITPC), 45, 65
Health Information Technology Standards Committee, 45
Health Insurance Portability and Accountability Act of 1996 (HIPAA)
administrative simplification by, 256, 257, 266
data warehouses and, 135
Enforcement Rule, 289, 299, 317
healthcare transactions affected by, 265, 266
Omnibus Rule, 256, 258, 286, 292, 308, 366
overview, 365
on privacy. See Privacy Rule
on security. See Security Rule
training on, 288, 318, 373

Transactions Rule, 266
Health Level 7 (HL7)
Clinical Information Modeling Initiative, 53
Consolidated Clinical Data Architecture, 53
data segmentation standards, 65
EHR System Functional Model, 512
Fast Healthcare Interoperability Resources, 54, 143, 176, 278
Functional Profile Registry, 512
in genomics and personalized medicine, 151–52
LOINC adoption for clinical messages, 52
Records Management and Evidentiary Support Functional Profile, 512
Version 2 bar-delimited ASCII strings, 53
Health literacy, 71, 375–76, 384–86, 414, 416, 423
Health maintenance organizations (HMOs), 258, 461
Health outcomes, 67, 179, 364, 379, 385, 432, 452–53
Health plans
compliance with HIPAA, 257
confidential communications requests to, 274
as defined by HIPAA, 258
ethical concerns regarding, 443
information systems of, 33, 66
research use of patient information from, 167
standard identifiers for, 256
Health records. See also Electronic health records; Personal health records
documentation in, 148
legal, 156, 157, 285, 418
paper, 286, 415
Health Resources and Services Administration (HRSA), 171
Health trend analysis, 45, 46, 71
Healthy People 2030 (HHS), 438–39
HHS. See Health and Human Services Department
HI. See Health informatics
HIEs. See Health information exchanges
HIMSS. See Health Information and Management Systems Society
HIOs. See Health information exchange organizations
HIPAA. See Health Insurance Portability and Accountability Act of 1996
Hippocratic Oath, 364
Histograms, 216–18
Historical research, 170
History, as threat to internal validity, 177
HIT. See Health information technology
HITAC. See Health Information Technology Advisory Committee
HITECH Act. See Health Information Technology for Economic and Clinical Health Act of 2009
HITPC. See Health Information Technology Policy Committee
HIV/AIDS, 65, 71, 74, 152, 167
HL7. See Health Level 7
HMOs. See Health maintenance organizations
Holographic data input, 455

Home monitoring devices, 17, 30, 90, 364
Hospital Compare website, 146
Hospital Transparency Rule, 22
Hot sites, 327
HP Vertica Analysis Platform, 185
HRSA. See Health Resources and Services Administration
Human Genome Project, 455
Human organs-on-chips, 455
Human resource management systems (HRISs), 87, 97, 98, 103–5, 107
Human threats to security, 301
Huntington's disease (HD), 380
HVA. See Hazard vulnerability assessment
Hypotheses, 167, 168, 174, 176, 222–23

I

IBM, 178, 185, 188, 190, 192, 234, 454
ICC. See Integrated Care Collaborative
ICD. See International Classification of Diseases
ICD-10-CM (International Classification of Diseases, Tenth Revision, Clinical Modification)
applications of, 46, 464
chapter titles, 543–44
code format in, 46–47
data table records and, 114
ICD-11 accommodations for, 48
repetition of injury patterns in, 49
transition from ICD-9-CM to, 47
ICD-10-PCS (International Classification of Diseases, Tenth Revision, Procedure Coding System), 46–48, 114, 464, 544
ICT. See Information and communication technology
Identity theft. See Medical identity theft
Identity Theft Prevention Program, 322–23
IDSs. See Integrated delivery systems
IGAM. See Information Governance Adoption Model
IGPHC. See Information Governance Principles for Healthcare
Implementation of health information systems, 85–109
adaptable solutions and, 107–8
adoption phase of, 90, 92–93
balancing cost, quality, and time in, 102–3
case studies, 103–4, 108–9
core skill requirements for, 11
costs of, 96, 102–3, 106–7
documentation of, 100, 108
organizational leadership for, 89–92
processes involved in, 94–103
stakeholders in, 86–87, 90–94, 103
systems development life cycle and, 95–102
team-based, 87–90, 93, 97–101, 107
technology considerations, 104–8

unintended consequences of, 154
Implementation phase of SDLC, 97–102, 106
Implementation specifications, 299, 300, 303–15
Implicit biases, 6
Import function, 190
Imputation, 215–16, 237
Incidental use of PHI, 268–69, 273
Independent variables, 173, 174, 177, 180, 183, 223–24
Infant mortality rates, 46
Inferential statistics, 174, 176, 180–82, 222–23
Infodemics and infodemic management, 444–46
Informatics. See also Health informatics; Public health informatics
biomedical, 368, 432, 475–76
clinical, 475, 476
medical, 413, 475
nursing, 475–76
Informatics for Integrating Biology and the Bedside (i2b2), 136–37
Information. See also Access to information; Health information; Knowledge
availability of, 28, 32, 33
consumer-generated, 16
in DIKW pyramid, 9–11
genetic, 257, 286, 308, 366, 379
genomic, 152, 459
just-in-time, 155
release of, 65, 267, 375, 383
requests for, 96
Information and communication technology (ICT), 398–400, 405, 407, 408
Information asset protection, 325–28
Information blocking, 20, 56, 66, 276–83
Information governance, 135, 354–57
Information Governance Adoption Model (IGAM), 356
Information Governance Principles for Healthcare (IGPHC), 355–56
Information systems. See also Health information systems
clinical, 16, 18, 20, 27, 30
community health management, 56–57
emergency department, 19
of health plans, 33, 66
laboratory, 17, 24, 67
operating room, 19
radiology, 8, 17, 24
Information technology. See Health information technology
Informed consent, 172, 364, 385
Inpatient care settings, 27–30, 32
In silico clinical trials, 455
Instagram, 384
Institute of Medicine. See National Academy of Medicine
Institutional review boards (IRBs), 167, 171–72
Instrumentation, as threat to internal validity, 177
Integrated Care Collaborative (ICC), 59

Integrated delivery systems (IDSs), 59
Integrated mobile devices, 18
Integration
of data, 11–12, 17, 151, 225, 465
of digital health technology, 399–401, 418–19
of EHR systems, 21, 23, 24, 30–31
of ethics policies, 374
of governance strategies, 354
of home monitoring devices, 90
of knowledge into practice, 10
of telemedicine services, 31
Integration testing phase, 98
Integrity
of data. See Data integrity
in information governance, 356
of statistical analysis, 363
Integrity controls implementation specification, 315
Integrity standard, 314
Interfaces
APIs, 56, 277–78, 407
EHRs and, 17, 18, 23, 24, 30
interactive, 189
for population health management, 70
for technical systems, 88
user interfaces, 17, 148, 156
Internal stakeholders, 87
Internal validity, 177
International Classification of Diseases (ICD). See also ICD-10-CM
computer-assisted coding and, 145
COVID-19 pandemic and, 47–48
health trend analysis and, 45, 46, 71
ICD-9-CM, 47
ICD-10-PCS, 46–48, 114, 464, 544
ICD-11, 48–49
in mortality and morbidity reporting, 46, 48, 206
ontology and, 48, 136
Internet computing, 18, 365, 412
Internet of Things (IoT), 399, 414, 418–19
Interoperability
in ACOs, 461
case study, 74–75
of CDSSs, 143
defined, 69
EHRs and, 17, 20–21, 23, 52–55
of health data, 59
health information exchanges and, 55, 56, 59
importance of, 4
MU program on, 20, 63
PHR platforms and, 17, 382

for population health, 70–71
Promoting Interoperability Programs, 56, 58, 70
for public health, 71–74
semantic, 53
Interval data, 174, 210
Interviews, as data collection method, 168, 175
Intimate space, 412
IoT. See Internet of Things
IRBs. See Institutional review boards
Issues Log workbook, 545–46
i2b2. See Informatics for Integrating Biology and the Bedside

J

Jobs. See Workforce
Join queries, 129, 130
Joint Commission (JC), 150, 377–78
JSON format, 54
Justice Department, 317
Justice, ethical principle of, 365, 376–77, 403, 471
Just-in-time information, 155

K

K-means clustering, 243
K-nearest neighbor algorithm, 242
Knowledge. See also Information
of business operations, 101
clinical, 467, 468, 475
data transformed into, 2, 5, 9–12, 95
in DIKW pyramid, 9–11
integration into practice, 10
Knowledge life cycle, 3
Knowledge management (KM), 143–44, 154, 192

L

Laboratory data, 49, 51–52, 73, 212, 439
Laboratory information systems (LISs), 17, 24, 67, 212
Laggards, 92–93
Leadership
of CDSSs, 150
ethical, 374–76
of implementation teams, 88, 93, 97
organizational, 89–92, 96
in public health informatics, 437
in risk mitigation, 303
Leapfrog Group, 146
Legal considerations
ethics and, 365–66
preemption of state law, 285
privacy issues and, 284–87

Legal counsel, 156, 260, 510
Legal health records, 156, 157, 285, 418
Legal holds, 285–86
LGBTQIA+ communities, 149–50
Life expectancy, 46, 184, 206
Likert rating scales, 210
Limited data sets, 269, 273
Linearizations, 48, 49
Linear regression, 181, 222, 224, 241, 246
Line charts, 187, 220
LinkedIn, 384
LISs. See Laboratory information systems
Literacy
    computer, 416
    data, 12
    health, 71, 375–76, 384–86, 414, 416, 423
    for treatment plan comprehension, 149
Logical data models, 131–32
Logical Observation Identifiers Names and Codes (LOINC), 51–52, 67, 152, 464
Login monitoring, 305–6
Logistic regression, 178, 185, 224, 242
Long-term care (LTC) settings, 32–33, 151

M

Machine learning (ML), 240–51
    algorithms used in, 471
    challenges related to, 6, 251
    in data analysis, 5, 11–12, 227
    decision support systems and, 68, 240
    deep learning, 7, 180, 207, 471
    for disease prediction, 244
    ethical concerns regarding, 471
    feature engineering and, 237–39
    performance evaluation of, 245–50
    predictive modeling and, 462
    Python statistical software and, 180
    reinforcement, 243
    supervised, 241–42
    unsupervised, 242–43
MAE. See Mean absolute error
Mainframe computing, 18
Maintenance records, 310
Majority adopters, 92
Malicious software (malware), 305–6, 402, 457, 470, 473
Management of data. See Data management
Mango Health, 454
Man-made disasters. See Disaster planning
Mann-Whitney test, 223
Many-to-many cardinality, 118, 119

Mass-casualty events, 73
Master patient index (MPI), 28, 58, 511
Maturation, as threat to internal validity, 177
Maturity levels of information governance, 356–57
MDS. See Minimum Data Set
Mean, 181, 212, 215, 222, 237
Mean absolute error (MAE), 247
Meaningful Use (MU) program
    EHRs and, 7, 20–21, 28, 55, 382
    health information exchanges and, 55–57, 63
    on interoperability, 20, 63
    LOINC and, 52
    renaming of, 56, 58
    RxNorm and, 52
    SNOMED CT and, 50
    tethered PHRs and, 20, 108, 415
Mean squared error (MSE), 246–47
Media controls standard, 311–12
Median, 181, 212, 215, 221
Media reuse implementation specification, 312
Medicaid, 62, 104, 258. See also Centers for Medicare and Medicaid Services
Medical devices
    data collection from, 24, 205
    FDA regulation of, 413
    innovation in, 418
    monitoring function of, 16
    sentinel events and, 378
    tethered, 90
    wearable, 7
Medical directors, 88, 92
Medical errors, 21, 346, 378, 381, 460
Medical ethics, 257, 362, 403
Medical homes, 459–60
Medical identity theft, 299, 321–25, 547–48
Medical imaging data, 7, 29–30, 207, 457
Medical informatics, 413, 475
Medical Information Mart for Intensive Care (MIMIC-III), 227
Medical records
    accounting of disclosures, 325
    data exchange of, 53
    electronic, 18–20, 28–31, 114, 376, 417
    patient copies of, 148
Medicare. See also Centers for Medicare and Medicaid Services
    ACOs and, 70, 460, 461
    claims processing by, 18, 20
    digital health coverage by, 403
    as HIPAA-defined health plan, 258
    Merit-Based Incentive Payment Systems, 70, 114
    reimbursement for services, 23, 70

Shared Savings Program, 461
Medications. See also E-prescribing
administration of, 28–30
compliance with, 147–48, 454
interactions involving, 53, 157, 388, 452
mHealth and, 411
over-the-counter, 73
patient education on, 71
patient safety related to, 29, 71
in precision medicine, 466
RxNorm codes for, 52–53, 464
sentinel events involving, 378
Medicine. See also Healthcare; Healthcare providers; Telemedicine
evidence-based, 29, 455, 460, 468
personalized, 151–52, 379, 398, 455, 465
precision, 225, 459, 464–65
MediSpan, 53
MedSpacy, 185–86
MedTagger, 185
Mental health care. See Behavioral healthcare
Merit-Based Incentive Payment Systems (MIPS), 70, 114
Metadata
data consolidation and, 458
data dictionaries and, 126
in data governance framework, 344
defined, 191, 208, 285
litigation and, 156, 285–87
management of, 192, 459
statistical analysis of, 146
MetaMap, 185
Metaverse, 455–56
Method, in object-oriented databases, 120
mHealth (mobile health), 401–2, 410–13, 419
Microsoft Access, 120, 133, 189–90, 192
Microsoft Analysis Services, 185
Microsoft Excel
as data collection tool, 175
flat file format in, 116, 117
importing data to, 190, 191
missing data in, 215
for reporting data, 191, 192
for statistical analysis, 177–78, 181
system building and development with, 97
for visualization of data, 187
Microsoft SQL Server, 120, 133, 137, 190, 192
Million Hearts Program, 467
MIMIC-III. See Medical Information Mart for Intensive Care
Minimum Data Set (MDS), 32
Minimum necessary requirement, 167, 266

Mining. See Data mining
MIPS. See Merit-Based Incentive Payment Systems
Missing data, 205, 211, 214–16, 235, 237, 458
Mitigation
    of medical identity theft, 325
    of public health crises, 72
    of risk, 290, 301, 303
    of security incidents, 306
    for wrongful PHI disclosures, 288
Mixed methods research, 169
ML. See Machine learning
Mobile computing add-ons, 107
Mobile health. See mHealth
Mobile technology
    growth in use of, 312
    health-related, 411–12
    implementation process for, 105–6
    integrated mobile devices, 18
    patient engagement through, 151
    reference sources accessed with, 388
    smartphones, 4, 7, 23, 365, 377, 411–12
Mode, 181
Monitoring
    of care outcomes, 70
    in data governance, 344
    for duplicate orders, 25
    of health conditions, 16, 28, 30
    of health information exchanges, 57
    home devices for, 17, 30, 90, 364
    of ICD codes, 47, 48
    for medical identity theft, 323, 324
    of quality of care, 73
    remote, 151, 379, 403, 408, 411, 413
    of system logins, 305–6
Morals, defined, 363. See also Ethics
Morbidity rates, 45, 48, 49, 71
Mortality rates, 45–46, 48–49, 71, 206, 238, 378
MPI. See Master patient index
MSE. See Mean squared error
Multifactor authentication, 314
Multifunctional radiology, 455
Multiple regression, 184
Multivariate regression splines, 185
Multum, 464
MU program. See Meaningful Use program
My HealtheVet, 416–17
MySQL, 120, 133

N

Naïve Bayes algorithm, 242
Named entity recognition (NER), 242
Narrative data, 24
Narrative response format, 512
National Academy of Medicine (NAM), 20, 25–26, 149–50, 440
National Academy of Sciences, Engineering, and Medicine (NASEM), 12
National Cancer Institute (NCI), 466
National Center for Biotechnology Information (NCBI), 459
National Center for HIV, Viral Hepatitis, STD, and TB Prevention Program Collaboration and Service Integration (NCHHSTP PCSI), 439
National Credit Union Administration, 322
National Drug Code (NDC), 136
National Drug File Reference Terminology, 53
National Health and Nutrition Examination Survey (NHANES), 221–22, 227
National Health Information Network, 62
National Institute of Standards and Technology (NIST), 300, 301, 319, 325–26, 472, 512
National Institutes of Health (NIH), 136, 467
National Library of Medicine (NLM), 50, 52, 150, 185, 203
National Notifiable Diseases Surveillance System (NNDSS), 72, 442, 539–41
National Vital Statistics System (NVSS), 46, 206
National Vulnerability Database, 301
Natural disasters. See Disaster planning
Natural experiments, 171
Natural language processing (NLP), 145, 152–53, 185–86, 207
Natural threats to security, 301
NCBI. See National Center for Biotechnology Information
NCHHSTP PCSI. See National Center for HIV, Viral Hepatitis, STD, and TB Prevention Program Collaboration and Service Integration
NCI. See National Cancer Institute
NDC. See National Drug Code
NER. See Named entity recognition
Neural networks algorithm, 242
Neuro-link, 453
Neurotrack, 453
Next Generation Accountable Care Organizations (NGACOs), 461
NHANES. See National Health and Nutrition Examination Survey
NIH. See National Institutes of Health
NIST. See National Institute of Standards and Technology
NLM. See National Library of Medicine
NLP. See Natural language processing
NNDSS. See National Notifiable Diseases Surveillance System
No-consent model, 60, 61
Nominal data, 174, 209–10, 221
Nonexperimental research, 170
Nonmaleficence, 365, 372, 376, 383–84, 403, 471
Nontechnical security vulnerabilities, 300–301
Nontextual unstructured data, 207
Normalized databases, 118–25, 237, 353, 464, 503–7
Notice of Privacy Practices (NPP), 61, 271, 275, 287, 288

Not Only Standard Query Language (NoSQL) databases, 5, 115
Null hypotheses, 174, 223
Numerical data, 168, 210, 215, 216
Nursing informatics, 475–76
NVSS. See National Vital Statistics System

O

Object-oriented databases, 5, 120
Observation, as data collection method, 175
Office for Civil Rights (OCR), 257, 263, 275, 289, 291, 298, 316–18
Office for Human Research Protections (OHRP), 171–72
Office of Disease Prevention and Health Promotion (HHS), 433, 438
Office of Minority Health (HHS), 150
Office of the Inspector General (HHS), 278
Office of the National Coordinator for Health Information Technology (ONC)
    Cures Act Final Rule, 23
    on data governance management, 353–54
    EHR adoption rate estimates, 28
    on electronic health information, 276
    Federal Healthcare IT Strategic Plan, 472
    health information exchanges and, 56, 57, 65–67, 152
    HITAC recommendations to, 45
    on information blocking, 278
    National Health Information Network established by, 62
    on patient engagement, 150
    Safety Assurance Factors for EHR Resilience, 157
    Standards and Certification website, 513
    Standards and Interoperability Framework, 149
    Value Set Authority Center and, 203
OHRP. See Office for Human Research Protections
Omnibus Rule (HIPAA), 256, 258, 286, 292, 308, 366
ONC. See Office of the National Coordinator for Health Information Technology
One-to-many cardinality, 118, 119
One-to-one cardinality, 118, 119
One-way between-subjects ANOVA, 183–84
Online bill payment, 31, 90, 109
On-Time Quality Improvement program, 151
Ontology, 48, 50, 136
OpenCDS consortium, 143
Open-source data science tools, 12
Operating room information systems (ORISs), 19
Operating systems, 105, 179
Operational costs, 107
Operational data, 114, 134
Operational layer of data governance, 354
Operations, as defined by HIPAA, 268
Operations, healthcare, 268, 269, 272, 273, 364
Opt-in model, 60–62
Opt-in with restrictions model, 60, 62

Opt-out model, 60, 61
Opt-out with exceptions model, 60, 61
Oracle, 115, 120, 133, 137, 185, 190
Order entry/management, 25–29, 142, 154–55
Order sets, 97, 107, 142, 156–57
Ordinal data, 174, 210, 221
Organizations. See also Healthcare organizations
    decision-making in, 88–90, 115, 142
    governance of, 88–90, 97, 99, 106–8
    HIO structures, 58–59
    information systems within, 146, 328
    leadership in, 89–92, 96
    recommendations for action by, 192–93
    reputation of, 64, 326, 383
    strategies of, 87, 89, 96
ORISs. See Operating room information systems
Outliers, 181, 187, 216, 237, 242, 363
Over-the-counter medications, 73

P

PACS. See Picture archiving and communication system
Pandemic. See COVID-19 pandemic
Paper charts, 30, 98, 166, 379, 456
Paper records, 270, 286, 288, 415, 416
Password management, 305–6, 473
Patient-centered care, 20, 386, 467, 469
Patient-centered medical homes (PCMHs), 459–60
Patient education
    CDSSs and, 149
    EHR resources for, 21–22
    on medications, 71
    mHealth and, 413
    self-education, 463
    on self-management of disease, 466
    by VHA, 462
Patient engagement, 146–51, 154, 452, 461, 467
Patient-generated health data (PGHD), 383, 402, 414, 418
Patient identifiers in EHRs, 24
Patient portals
    as administrative DSSs, 145
    Blue Button, 146–48
    consumer perceptions of, 65–66
    data collection from, 463
    data mining of, 143
    e-health and, 414
    information blocking and, 278
    PHI access through, 381, 382
    public health functions of, 74
    tethered PHRs and, 33

training on use of, 146
Patient safety
decision support systems and, 142, 146, 148–51, 157
EHRs in support of, 25
IOM reports on, 20
medication-related, 29, 71
Patient support function of EHRs, 25, 26
Payment. See Reimbursement for services
PCA. See Principal component analysis
PCMHs. See Patient-centered medical homes
PDB. See Protein Data Bank
Pearson's correlation test, 180–82
Penalties for HIPAA noncompliance, 289, 317–19
Performance-based programs, 21
Personal computers, 18, 99, 128, 382, 387, 412, 473
Personal disaster plans, 328
Personal health records (PHRs). See also Tethered PHRs
consumer perceptions of, 66
digital technology and, 415–18
ethical concerns regarding, 382–83
implementation of, 87, 89–90, 416–17
interoperability of platforms, 17, 382
proliferation of, 18, 415
stand-alone, 33, 382, 416
Personalized medicine, 151–52, 379, 398, 455, 465
Person or entity authentication standard, 314
PGHD. See Patient-generated health data
Pharmacogenomics, 225, 464–65
Pharmacy databases, 136
PHI. See Protected health information; Public health informatics
PHIN. See Public Health Information Network
Phishing, 299, 332, 402
PHRs. See Personal health records
Physical data models, 131–33
Physical safeguards, 260, 288, 299, 309–12, 318
Physician Quality Program Reporting System (PQRS), 70
Physicians. See Healthcare providers
Picture archiving and communication system (PACS), 18, 24, 28, 29, 207, 457
Pie charts, 187, 219–20
PivotCharts, 187, 191, 192
Placebo, in experimental research, 169
Planning phase of SDLC, 92–93, 96, 99, 100, 103, 107
Point-of-care (POC) systems, 28, 29
Point-of-care testing (POCT), 463
Policy development, 90, 318, 372, 432–33, 435, 476
Population health
assessment of, 432
decision-making for, 463–64
defined, 69, 443–44, 463

interoperability for, 70–71
management of, 26, 70, 108
promotion of, 442
Population health informatics, 443–44
Population, in research studies, 174
Population-level quality monitoring, 73
Portals. See Patient portals
Positive predictive value (PPV), 248
Post-conversion support, 100–101
Post-coordination in ICD-11, 49
PostgreSQL, 120, 133
PQRS. See Physician Quality Program Reporting System
Precision, as evaluation measure, 248
Precision medicine, 225, 459, 464–65
Precision of data, 348
Predictive analysis, 184
Predictive analytics, 234, 462
Predictive modeling, 462
Preemption of state law, 285
Preprocessing of data, 211, 236–37
Prescribing. See E-prescribing
Prescription drugs. See Medications
Prescriptive analytics, 235, 462
Presentation of data, 192, 202, 203, 226, 452
Preventive healthcare, 71, 142, 244, 464–66
Price transparency, 20, 22, 54
Primary data use, 25, 166, 408
Primary keys, 117–19, 126, 191, 503
Principal component analysis (PCA), 239
Privacy issues, 255–92. See also Privacy Rule (HIPAA); Protected health information
in aggregation of data, 71
in big data analytics, 225–26
CHMISs and, 57
EHR systems and, 7, 23, 24, 288
electronic communications and, 146
electronic health information and, 66
health information exchanges and, 57, 58
health information systems and, 5
legal considerations, 284–87
in machine learning, 251
Security Rule on, 316
Privacy officers, 288, 303
Privacy Rule (HIPAA). See also Covered entities; Protected health information
on breaches, 257, 289–92
on business associates, 257–60, 262, 273, 289–92
on documentation and record retention, 288
filing complaints about violations of, 275, 288
on individual rights to information, 270–75
Notice of Privacy Practices, 61, 271, 275, 287, 288

OCR oversight and enforcement of, 257, 289, 291
penalties for noncompliance, 289, 318–19
on research use of patient information, 167
on safeguards, 260, 269, 288
waiver of rights under, 288
workforce as defined by, 258
Private HIOs, 59
Procedural coding systems
CPT, 50, 101, 114, 117, 136, 145
HCPCS, 50
ICD-10-PCS, 46–48, 114, 464, 544
Process mapping, 462
Program implementation, 87–89
Program managers, 88–90, 100, 101
Project Health Design, 417
Project implementation, 87–90
Project managers, 88, 90, 97, 100, 101
Promoting Interoperability Programs, 56, 58, 70
Protected health information (PHI), 261–75. See also Disclosure of protected health information; Electronic protected health information; Privacy Rule (HIPAA)
access to, 270–71
amendment requests, 271–72
best practices for protection of, 473
breaches related to, 289–92, 298–99, 319–20
confidential communications requests, 274
as defined by HIPAA, 257, 261–62
deidentified, 135–36, 167, 172, 227, 263–66
ethical concerns regarding, 379
genetic information as, 257, 286, 308, 366, 379
limited data sets, 269, 273
minimum necessary requirement and, 167, 266
researcher's use of, 167, 263
state laws on, 285
use as defined by HIPAA, 266–67
Protection principle for information governance, 356
Protein Data Bank (PDB), 470
Provisioning of care, 71, 142, 364, 375, 376, 461
Psychotherapy notes, 270, 276
Public health
alert systems for, 73–74
clinical care and, 73
core functions and essential services of, 432–33
data on, 44–46, 56, 72–73, 433, 439, 440
defined, 69, 71, 432
health information exchanges and, 58, 59, 72–74, 440
infrastructure for, 438–39
interoperability for, 71–74
investigations involving, 73, 456
management of, 414

surveillance of, 63, 72–73, 285, 440–42
Public health informatics (PHI), 431–46
competencies for, 433–37
defined, 186, 432
globalization of, 444–46
Healthy People 2030 on, 439
population health informatics vs., 443–44
real-world applications of, 442
role of data in, 439
surveillance and, 440–42
Public Health Information Network (PHIN), 440, 441
Public HIOs, 59
PubMed, 192
Python, 178–80, 185–86, 189, 215, 220–21
PyTorch, 180

Q

QPP. See Quality Payment Program
Qualitative data. See Categorical data
Qualitative research, 21–22, 154, 168–69
Quality improvement (QI), 32, 44, 50, 151, 460, 462–63
Quality of care
accountability and, 20, 70
in behavioral healthcare, 32
clinical outcomes assessment and, 171
CMS surveys on, 143
consumer demands for, 7
decision support systems and, 142–46, 148, 155–56
digital health technology and, 408, 409, 423
health information exchanges and, 64, 74
improvements in, 30, 204, 452, 461
population-level monitoring of, 73
societal concerns regarding, 18
technology platforms and, 16, 21
Quality of data. See Data quality
Quality Payment Program (QPP), 114
Quantitative data. See Continuous data
Quantitative research, 168, 169, 173, 193
Quasi-experimental research, 170–71

R

Radio frequency identification (RFID), 29, 103–5
Radiology information systems (RISs), 8, 17, 24
Random forest model, 185
Random sampling, 169
Range, 181, 212
Ranked data. See Ordinal data
Ransomware, 332, 457, 471
RapidMiner, 185

Ratio data, 174, 210
Raw data, 114, 203–4, 211, 225, 235–37
Realizing Equity, Access, and Community Health (REACH) model, 461
Real-time data, 354, 454
Real-time diagnostics, 455
Real-time interactivity. See Synchronous interactivity
Recall, as evaluation measure, 248–49
Receiver operating characteristic (ROC) curve, 250
Record locator services, 58
Records. See also Documentation; Health records; Medical records
in data tables, 114–16
maintenance, 310
management of, 251, 418, 511, 512
paper, 270, 286, 288, 415, 416
retention of, 288
Records Management and Evidentiary Support Functional Profile, 512
Recovery planning. See Disaster recovery planning
Recovery point objectives (RPOs), 326
Recovery time objectives (RTOs), 326
REDCap. See Research Electronic Data Capture
Red Flag Rules, 322–23
Redisclosure. See Secondary release of information
Redundancy of data, 73, 117, 118, 458
Regenstrief Institute, 51, 52
Regional Extension Centers (RECs), 467
Regional Health Information Organizations (RHIOs), 57
Registered Health Information Administrators (RHIAs), 476
Registered Health Information Technicians (RHITs), 476
Regression
linear, 181, 222, 224, 241, 246
logistic, 178, 185, 224, 242
missing data and, 216, 237
multiple, 184
statistical, 177
sum of squares, 246
Rehabilitation CIS systems, 30
Reidentification mechanisms, 263, 265
Reimbursement for services
digital health and, 402, 403, 422, 423
disclosure of PHI for, 268
health data for, 4, 43, 377
Medicare program and, 23, 70
Reinforcement learning, 243
Relational databases
adaptable solutions and, 107
cardinality in, 118, 119
description of, 115, 117
normalized, 118–25, 503–7
prevalence of use, 5, 18

software for management of, 120, 133
SQL and, 126, 128–30, 190
Release of information (ROI), 65, 267, 375, 383
Relevance of data, 349
Reliability, 176, 177, 463
Remote monitoring, 151, 379, 403, 408, 411, 413
Reporting
of breaches, 290–91, 319–20
of data, 44–45, 56, 72–73, 191–93, 204
in data governance, 344
on Identity Theft Prevention Program, 323
in implementation phase of SDLC, 98–99
of mortality and morbidity statistics, 46, 48, 206
self-reporting, 152, 175
training in, 98, 108
Reputation, 64, 326, 383, 514
Requests for information (RFIs)
comparison with RFPs, 509–10
identifying prospects, 510
in planning phase of SDLC, 96
response format for, 512
template for, 514, 517–37
Requests for proposals (RFPs)
comparison with RFIs, 509–10
developing and disseminating, 510–11
finalizing and submitting, 514
identifying prospects, 510
importance of process, 513
in planning phase of SDLC, 96
response formats for, 511–12
template for, 514, 517–37
value of HIM involvement with, 511
vendor selection using, 514–15
Required implementation specifications, 299, 300, 303, 305–6, 311–13
Research, 165–93
analysis of data in, 176–86
basic vs. applied, 167
clinical outcomes assessment, 171–72
correlational, 170
on decision support systems, 150–54
defined, 166
descriptive, 170
design principles, 167–72
EHRs for, 193
epidemiological, 186
experimental, 169–70
historical, 170
institutional review board approval of, 167, 171–72
mixed methods, 169

PHI used for, 167, 263
qualitative, 21–22, 154, 168–69
quantitative, 168, 169, 173, 193
quasi-experimental, 170–71
reporting of data in, 191–93
retrieval of data in, 189–90
retrospective (case control), 170
steps in, 173–76
visualization of data in, 180, 187–89, 192
Research Electronic Data Capture (REDCap), 175–76
Research questions, 167, 168, 173–74
Resolution agreements, 289, 318–19
RESTful web services, 54
Restriction requests, 274–75
Results management, 25, 26
Retaliation against complaints, 288
Retention principle for information governance, 356
Retrieval of data, 189–90
Retrospective research, 170
Return on investment (ROI), 67, 70, 423
Revenue codes, 119, 124, 503, 507–8
Revisions procedures, 307
RFID. See Radio frequency identification
RFIs. See Requests for information
RFPs. See Requests for proposals
RHIAs. See Registered Health Information Administrators
RHIOs. See Regional Health Information Organizations
RHITs. See Registered Health Information Technicians
Risk
analysis of, 300–303, 320, 324
assessment of, 289–90, 300, 302, 320, 325–26, 329, 387
management of, 99–100, 179, 303, 377, 387–88, 472
mitigation of, 290, 301, 303
RISs. See Radiology information systems
ROC curve. See Receiver operating characteristic curve
Rogers Adoption Curve, 92
ROI. See Release of information; Return on investment
RPOs. See Recovery point objectives
R software, 178, 179, 185, 186, 188, 215, 220
R-squared (R^2), 246
RStudio, 179, 180
RTOs. See Recovery time objectives
RxNorm, 52–53, 464

S

Safeguards
administrative, 260, 288, 299, 302–8
physical, 260, 288, 299, 309–12, 318
Privacy Rule on, 260, 269, 288

Security Rule on, 299, 302–15
technical, 260, 288, 299, 312–15, 318
Safe harbor method, 263–65
Safety. See Patient safety; Security issues
Safety Assurance Factors for EHR Resilience (SAFER), 157
Samples and sampling methods, 169, 174, 175
Sanctioning rules in ICD-11, 49
Sanctions policies, 303
SANs. See Storage Area Networks
SAS. See Statistical Analysis Software
Scalability testing phase, 98
Scatterplots, 183, 188, 222, 224
SDLC. See Systems development life cycle
Seamless care, 468–69
Secondary data use, 25, 44, 50, 166, 204, 206, 263
Secondary release of information, 65, 267, 375
Second normalized form, 118, 119, 503
Security awareness and training standard, 305–6
Security incident, defined, 299
Security incident procedures standard, 306
Security issues, 297–333. See also Security Rule (HIPAA)
in aggregation of data, 71
in analysis phase of SDLC, 96
in big data analytics, 225–26
case studies, 332–33
cybersecurity, 263, 469, 471–73
data level security, 99
data warehouses and, 135
for digital health technology, 402
disaster planning and, 325–32
EHR systems and, 23, 24
electronic health information and, 66, 262
hackers, 299, 301, 325–26, 332, 381, 402, 472
health information exchanges and, 57, 58, 152
health information systems and, 5, 314
in machine learning, 251
medical identity theft, 299, 321–25, 547–48
phishing, 299, 332, 402
smishing, 402
threats, 301, 302, 326–27
vulnerabilities, 299–302, 320, 332, 381
Security management process standard, 302–3
Security officers, 303
Security reminders, 305–6
Security risk analysis, 300–303, 320, 324
Security Rule (HIPAA). See also Covered entities
on availability of data, 306, 316
on business associates, 299–302, 307–8, 320
on confidentiality of data, 316

on ePHI, 262–63, 299–309, 311–15
implementation specifications of, 299, 300, 303–15
on information asset protection, 325–28
on integrity of data, 314, 316
OCR oversight and enforcement of, 316–18
penalties for noncompliance, 317–19
on safeguards, 299, 302–15
on security risk analysis, 300–303, 320
standards related to, 299, 300, 303–15
Segmentation of data, 65
Selection bias, 177
Select queries, 128–29
Self-management of disease, 466
Semantic gaps, 8
Semantic interoperability, 53
Sensitivity, as evaluation measure, 248–49
Sensors
digestible, 455
physiological, 5
for real-time data collection, 454
wearable, 8, 205, 215, 225
Sentinel events, 377–79
Services implementation, 106
Sign-off documents, 511
SmartCare concept, 419
Smartphones, 4, 7, 23, 365, 377, 411–12
Smishing, 402
SNOMED. See Systemized Nomenclature of Medicine
SNOMED CT. See Systemized Nomenclature of Medicine – Clinical Terms
Social determinants of health, 32–33, 239, 444
Social media, 65, 146, 384, 386. See also specific platforms
Social networking, 376, 384–85
Software. See also specific programs
access controls for, 310, 313
antivirus, 365, 381, 472, 473
database management, 120, 133
data reporting, 191–92
implementation of, 104–5
interfaces between platforms, 88
malicious, 305–6, 402, 457, 470, 473
presentation, 192
statistical analysis, 176–80, 191, 215
updates and upgrades for, 105
voice recognition, 21
Software-as-a-service licensing, 108
Software solution configuration, 85
Solve the Outbreak, 456
Specificity, as evaluation measure, 249
Splitting data sets, 245, 246

Sponsors, of health system implementations, 87–90, 93, 94, 96–99, 101
Spread, measures of, 181
Spreadsheets, 114–17, 175, 178, 215, 351
SPSS. See Statistical Package for the Social Sciences
SQL. See Structured Query Language
Stakeholders
communication with, 93, 100, 103
in digital health technology, 398, 400–401, 421–23
in disaster planning, 329
EHR systems and, 20–21, 33
in health informatics, 16
in health information exchanges, 64–65
in implementation of health information systems, 86–87, 90–94, 103
management of, 105
in RFP process, 510, 513
in SDLC, 96, 100, 102
types of, 87, 92–93
Stand-alone PHRs, 33, 382, 416
Standard deviation, 181, 212, 237
Standards
for CDSSs, 149
for code sets, 256
data, 44, 176, 202–3, 350, 433, 440
for EHR technologies, 23
for federal advisory committees, 45
in health information systems, 44
of practice, 366–74
of Security Rule, 299, 300, 303–15
Standards of Ethical Coding (AHIMA), 368
Stanford Chronic Disease Self-Management Program (CDSMP), 466
Stanford Medicine Research Data Repository (STARR), 136–37
STAR HIE. See Strengthening the Technical Advancement and Readiness of Public Health Agencies via Health Information Exchange
Stata software, 179, 190
State Department, 325
State Health Information Exchange Cooperative Agreement Program, 56, 62
State law, preemption of, 285
Statistical analysis, 176–86, 191, 215, 363
Statistical Analysis Software (SAS), 179, 185, 190
Statistical modeling, 223–24
Statistical Package for the Social Sciences (SPSS), 178–85, 188, 190–91, 215, 220
Statistical regression, 177
Statistical tests, 168, 174, 176, 178, 180–82, 223
Statistical text mining (STM), 185
Stewardship, 345, 353–54, 401
Storage Area Networks (SANs), 457
Storage of data, 311, 312, 316, 318–19, 457–58
Store-and-forward interactivity. See Asynchronous interactivity
Strategic layer of data governance, 354

Strengthening the Technical Advancement and Readiness of Public Health Agencies via Health Information Exchange (STAR HIE), 56
Structured data
advantages and disadvantages, 206
characteristics of, 8
in clinical documentation, 24
consolidation of, 458
data entry of, 24, 50, 205
in decision-making, 146
from EHRs, 456
examples of, 207, 235
integrity of, 342
levels of measurement for, 174
Structured Query Language (SQL), 126, 128–30, 190
Summary of care documents, 58
Summary statistics. See Descriptive statistics
Sum of squares regression, 246
Superusers, 101, 108
Supervised learning, 241–42
Support vector machines (SVMs), 242
Surveys, 168, 169, 175, 178, 193, 209, 215
Sustainment phase of SDLC, 101–2
Synchronous interactivity, 405–6, 408, 411
System building and development, 97
Systemized Nomenclature of Medicine (SNOMED), 48, 152
Systemized Nomenclature of Medicine – Clinical Terms (SNOMED CT), 50–51, 136, 464
Systems development life cycle (SDLC), 95–102
analysis phase of, 96
holistic nature of, 95–96
implementation phase of, 97–102, 106
planning phase of, 92–93, 96, 99, 100, 103, 107
stakeholders in, 96, 100, 102
sustainment phase of, 101–2
utilization of, 105

T

Tables. See Data tables
Tactical layer of data governance, 354
Team-based implementation, 87–90, 93, 97–101, 107
Technical safeguards, 260, 288, 299, 312–15, 318
Technical security vulnerabilities, 300
Technical structure of HIEs, 58
Technology
adoption of, 90, 92–93
AI. See Artificial intelligence
cloud. See Cloud computing
deep learning, 7, 180, 207
digital. See Digital health technology

emerging, 453–59, 469–73
HIT. See Health information technology
ICT, 398–400, 405, 407, 408
internet computing, 18, 365
mainframe computing, 18
ML. See Machine learning
mobile. See Mobile technology
natural language processing, 145, 152–53, 185–86, 207
personal computers, 18, 99, 128, 382, 387, 412, 473
social media, 65, 146, 384, 386
Technology platforms. See also Hardware; Software; specific platforms
diversity of, 104–6
efficiency of, 17, 20–22, 25, 26
quality of care and, 16, 21
sustainable use of, 90
updates and upgrades to, 105
value-added functions of, 16
Telehealth
advantages and disadvantages, 408–9
AMA playbook on, 401–2
beneficence and, 364
case studies, 332–33, 409–10
in COVID-19 pandemic, 333
for CRT, 409–10
defined, 407
expansion of, 407–9
Telemedicine
defined, 407
emergence and growth of, 398, 408
implementation of systems for, 87, 88
integration of services, 31
mixed methods study of, 169
proliferation of, 18
seamless care through, 468
technology platforms for, 16
Telerehabilitation, 399, 403, 406
TensorFlow, 180
Terminologies. See also Clinical terminologies; Vocabularies
for databases, 114–15
in health information systems, 44
standardization of, 26
Testing
clinical, 16, 17
of contingency plans, 307
for COVID-19, 63
of decision support systems, 156
genetic, 286, 366, 379–80
of hypotheses, 168, 174, 176
in implementation phase of SDLC, 97, 98

insurance claims for, 18, 20
point-of-care, 463
statistical, 168, 174, 176, 178, 180–82, 223
as threat to internal validity, 177
in trend analysis, 183–84
Tethered PHRs
adoption of, 108
advantages of, 33, 35–36, 382
data entry into, 66
defined, 17, 33
example of, 34
functions of, 35–36
implementation of, 108–9
MU program and, 20, 108, 415
online bill payment through, 31
uploading data into EHRs from, 30
Texas Health Services Authority, 60
Textual unstructured data, 207
Theft. See Medical identity theft
Third normalized form, 118, 120
Threats to security, 301, 302, 326–27
3M Health Information Systems, 47
Timelines for health system implementations, 102–3
Timeliness of data, 349
Time series analysis, 184
TNs. See True negatives
TPO. See Treatment, payment and operations
TPs. See True positives
Training. See Education and training
Transactions in healthcare, 5, 258, 265, 266
Transactions Rule (HIPAA), 266
Transmission media, 262, 311
Transmission security standard, 315
Transparency
of access logs, 422
of AI algorithms, 453
in breaches, 291
in decision-making, 374
Hospital Transparency Rule, 22
in information governance, 356
of pricing, 20, 22, 54
Treasury Department, 322
Treatment, payment and operations (TPO), 268, 272–74
Treemaps, 220
Trend analysis, 183–84
True negatives (TNs), 247–49
True positives (TPs), 247–49
T-tests, 183, 223
21st Century Cures Act of 2016, 20, 23, 45, 56, 275–77

Twitter, 146, 220, 384

U

UIMA. See Unstructured Information Management Architecture
Ultrafast scans, 453
Unified Medical Language System (UMLS), 152, 185
Unified Metalanguage (UML), 53
Uniform Rules Relating to Discovery of Electronically Stored Information, 285
Unintended consequences, 154, 155, 545–46
Unique user identification methods, 313
Unit testing phase, 98, 101
Unstructured data
    characteristics of, 8
    in clinical documentation, 24
    consolidation of, 458
    data entry of, 24, 207
    digital health technology and, 401
    from EHRs, 456
    examples of, 207, 235
    organization of, 206–7
Unstructured Information Management Architecture (UIMA), 185
Unsupervised learning, 242–43
Use cases, 72–74, 135, 149, 242, 452
User interfaces, 17, 148, 156

V

Validity, 176, 177, 180
Value Based Modifier for PQRS and Medicare, 70
Value-based purchasing programs, 21–22, 114
Value-based technology implementation, 106–8
Value of data, 205, 354
Values, defined, 363. See also Ethics
Value Set Authority Center (VSAC), 203
Variables
    cause and effect for, 170
    in databases, 116–20
    data dictionaries for, 116
    dependent, 173, 174, 180, 183, 223–24
    independent, 173, 174, 177, 180, 183, 223–24
    in quantitative research, 168
Variance, 181
Variety of data, 179, 205, 224, 354
Velocity of data, 204, 205, 354
Vendors
    CHIN fees charged by, 57
    collaborating with, 89
    of EHR platforms, 21, 23, 31, 148
    of EMR platforms, 18
    identifying prospects, 510

negotiating with, 105
RFI/RFP process, 96, 509–15
in SDLC, 96–102
selection of, 381, 467, 513–15
for services implementation, 106
as stakeholders, 87
testing plans from, 98
training provided by, 99–100
Veracity of data, 205, 354
Veterans Health Administration (VHA)
ACO status of, 461–62
ambulatory care clinics, 185
Blue Button portal, 146–48
Common Rule and, 167
health information exchange efforts, 59, 73, 152
My HealtheVet, 416–17
National Drug File Reference Terminology, 53
Videoconferencing, 406, 408, 410
Virtual reality (VR), 455–56
Visualization dashboards, 98, 146, 189
Visualization of data, 180, 187–89, 192, 216–20, 226
Vital statistics, 46, 69, 206
Vocabularies. See also Terminologies
assignment of meaning through, 4
for clinical documentation, 24, 29
data consolidation and, 458
for e-prescribing medications, 52
in health information systems, 44
of Public Health Information Network, 440, 441
Voice recognition software, 21
Volume of data, 7, 11, 193, 205, 225, 354, 474–75
Volume testing phase, 98
VSAC. See Value Set Authority Center
Vulnerabilities, security-related, 299–302, 320, 332, 381
Vulnerable research subjects, 172

W

Waiver of rights under Privacy Rule, 288
Warehousing of data. See Data warehouses
Warm sites, 327
Wearable devices
advancements in, 454, 455
health data from, 4, 7
health services delivery via, 399
mHealth and, 411, 413
sensors, 8, 205, 215, 225
Whistleblowers, 376–77
White hat hackers, 381, 472
WHO. See World Health Organization

Wilcoxon test, 223
Wisdom, in DIKW pyramid, 9–11
Word cloud diagrams, 217, 219
Workforce
    as defined by HIPAA, 258
    in health informatics, 474–81
    security standard for, 303–5
    shortages in, 212, 474
Workstation security standard, 311
Workstation use standard, 311
World Health Organization (WHO). See also International Classification of Diseases
    Composite International Diagnostic Interview, 151
    COVID-19 dashboard, 188
    on digital health technology, 398, 400
    on infodemic management, 444, 445

Y

YouTube, 384